

Weak gravitational lensing analyses of clusters of galaxies

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Karianne Holhjem
aus
Oslo, Norwegen

Bonn (Februar, 2011)

Angefertigt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinische Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Peter Schneider
2. Gutachter: Prof. Dr. Thomas H. Reiprich

Tag der Promotion: 20 April 2011
Erscheinungsjahr: 2011

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online/ elektronisch publiziert.

to His glory

Contents

Introduction and summary	1
1 Cosmology	3
1.1 An expanding Universe	4
1.2 Cosmic evolution	5
1.2.1 The Universe on small scales	5
1.2.2 The evolution of the cosmic scale factor	6
1.3 General Relativity	7
1.3.1 Obtaining the Friedmann equation from General Relativity	7
1.3.2 Obtaining the equation of motion from General Relativity	8
1.4 Matter models of the Universe	9
1.5 Cosmological parameters	10
1.5.1 Future expansion of the Universe	11
1.6 Redshift in the context of cosmic expansion	13
1.7 Distance measurements in a dynamic spacetime	13
1.8 The concordance model, Λ CDM	15
1.8.1 Evidence of dark matter	15
1.8.2 The flatness problem	16
1.8.3 The horizon problem	17
1.8.4 Inflation	18
1.8.5 Observations confirming the theory	19
2 Gravitational lensing	21
2.1 Introduction	21
2.1.1 History	22
2.2 Lensing geometry	22
2.2.1 The deflection angle	23
2.2.2 The lens equation	24
2.2.3 Magnification and distortion	25
2.2.4 Relative lensing efficiency	27
2.3 Shear measurements using weak lensing	28
2.3.1 Shape estimation	29
2.3.2 PSF corrections using the KSB+ formalism	32

Contents

2.4	Lens models	34	
2.4.1	Axially symmetric lenses	34	
2.4.2	The singular isothermal sphere	36	
2.4.3	The Navarro, Frenk, & White profile	37	
2.5	Two-dimensional mass reconstruction	39	
3	Clusters of galaxies		41
3.1	Optical clusters	41	
3.1.1	History	43	
3.1.2	Morphological classification	43	
3.2	X-ray clusters	45	
3.2.1	Morphological classification	46	
3.3	Measuring the cluster mass	46	
3.3.1	Virial analysis	46	
3.3.2	X-ray	48	
3.4	Structure formation	49	
3.4.1	The early Universe	49	
3.4.2	Linear perturbation growth	50	
3.4.3	Nonlinear growth of structure	51	
4	Data reduction		53
4.1	Pre-Processing	53	
4.1.1	Bias	54	
4.1.2	Dark current	55	
4.1.3	Flatfield	55	
4.1.4	Fringing	57	
4.2	Image combination	57	
4.2.1	Sky subtraction	57	
4.2.2	Bad pixel mask	57	
4.2.3	Astrometry and relative photometry	58	
4.2.4	Coaddition	59	
4.3	Photometry	60	
4.3.1	Magnitudes	60	
4.3.2	Commonly used filters	62	
4.3.3	Photometric calibration	63	
4.4	Redshift determination	64	
4.4.1	Photometric redshifts	64	
4.4.2	Spectroscopic redshifts	64	
4.4.3	Spectroscopic vs. photometric redshift determination	65	
5	Weak lensing analysis of Abell 1351 and Abell 1995		67
5.1	Observations and data reduction	68	
5.1.1	Data acquisition	68	
5.1.2	Image processing	68	
5.2	Shear measurements	71	
5.3	Mass reconstruction	71	

5.3.1	Mass and galaxy density distributions	74
5.3.2	Lower mass peaks in the fields	76
5.3.3	Giant radio halo in Abell 1351	77
5.4	Modelling the lensing data	77
5.4.1	Cluster contamination and magnification depletion	78
5.4.2	Fitting the SIS and the NFW profiles	79
5.5	Discussion	80
5.5.1	The mass estimates	82
5.5.2	The concentration parameter	83
5.5.3	Centre position	85
5.5.4	The mass reconstructions	86
5.6	Conclusions	86
6	Lensing by high-z galaxy clusters using LBGs in the COSMOS survey	87
6.1	The data set	88
6.1.1	The Cosmic Evolution Survey	88
6.1.2	The cluster list	88
6.1.3	The shear catalogue	88
6.1.4	The Lyman-break catalogue	89
6.1.5	The photometric redshift catalogue	90
6.2	Lensing measurements	91
6.2.1	Inner cluster radius	91
6.3	Modelling the lensing signal from Lyman-break galaxies	92
6.4	Modelling the lensing signal from non- and photo- z galaxies	94
6.4.1	Modelling the redshift distribution for shear galaxies lacking photo- z estimates	95
6.4.2	Photo- z vs. nonphoto- z galaxies	97
6.4.3	Cluster galaxy contamination	98
6.5	Fitting the SIS profile to the joint LBG, non- and photo- z catalogue	98
6.6	Consistency checks	99
6.6.1	B modes	99
6.6.2	Weights	99
6.6.3	Cluster redshift re-calculation	99
6.7	Conclusions	99
7	Analysis of the M/L ratio for galaxy clusters in the COSMOS field	101
7.1	Introduction	101
7.2	Cluster samples	103
7.2.1	Optically selected clusters	103
7.2.2	X-ray selected clusters	104
7.2.3	Common clusters	105
7.2.4	Projected sky coverage	105
7.3	Measuring cluster properties: luminosity and richness	106
7.3.1	Calculating luminosities of individual objects	107
7.3.2	Richness	109
7.3.3	Luminosity	110

Contents

7.4	Measuring the cluster mass	112
7.5	Results and discussion: mass-to-light ratio and mass-richness relation	115
7.5.1	Correction for passive evolution	115
7.5.2	Luminosity vs. richness	115
7.5.3	Mass-to-light ratio	116
7.5.4	The mass-richness relation	123
7.5.5	Mass-to-light as a function of central richness	125
7.6	Results and discussion: comparing X-ray and optical measurements	126
7.7	Summary and conclusions	130
A	Cluster tables for Chapter 7	131
	Bibliography	139
	Acknowledgements	145
	Declaration	147
	Curriculum Vitae	149

List of Figures

- 1.1 Classification of cosmological models 12
- 1.2 Constraints from SN Ia, CMB, and BAO on Ω_Λ and Ω_m 20

- 2.1 Geometry of a gravitational lens system 23
- 2.2 The effects of convergence and shear on an intrinsically circular source 28
- 2.3 Atmospheric PSF effects 29
- 2.4 Relation between ellipticity and orientation for Gaussian ellipsoid objects 30

- 3.1 Hubble's tuning fork diagram for galaxies 43
- 3.2 The Rood-Sastry classification system for galaxy clusters 44

- 4.1 Image reduction and calibration frames 61
- 4.2 Optical, infrared, and narrowband filter transmittance curves 63

- 5.1 Size-magnitude and weighted ellipticity parameter diagram 70
- 5.2 Star ellipticities before and after PSF anisotropy corrections 72
- 5.3 Projected surface mass densities and cross components 73
- 5.4 Colour-magnitude diagrams 74
- 5.5 Mass reconstruction and S -statistics 75
- 5.6 Percentage of cluster galaxies in the faint galaxy catalogue 78
- 5.7 Reduced tangential shear and cross-component as a function of radius 80

- 6.1 Resulting θ_E as a function of starting radius in the SIS fit 92
- 6.2 Redshift distribution for LBGs 93
- 6.3 SIS fits to individual dropout catalogues 93
- 6.4 SIS fits to different background galaxy samples 96
- 6.5 Magnitude distribution for the photo- z and nonphoto- z shear catalogues 97
- 6.6 Cluster galaxy contamination in shear catalogue for stacked cluster sample 98

- 7.1 Colour-colour diagram from galaxy SED templates 107
- 7.2 Central richness vs. total richness 111
- 7.3 Mass and luminosity as a function of redshift 113
- 7.4 Luminosity-richness relation 116
- 7.5 Mass vs. luminosity for total sample 117

List of Figures

- 7.6 Mass vs. luminosity in redshift bins 118
- 7.7 M/L_B as a function of redshift 119
- 7.8 Mass vs. uncorrected luminosity in redshift bins 121
- 7.9 Mass-richness relation in redshift bins 124
- 7.10 Mass vs. luminosity with respect to central richness 125
- 7.11 Optical vs. X-ray luminosity in redshift bins 127
- 7.12 Lensing mass vs. X-ray luminosity in redshift bins 128
- 7.13 Lensing mass vs. X-ray temperature and X-ray mass 129
- 7.14 Optical luminosity vs. X-ray temperature 129

List of Tables

- 5.1 Results from fitting theoretical density profiles to measured shear values 81
- 5.2 Results from different starting radii for Abell 1351 84

- 6.1 Number of LBGs in samples of different selection criteria, together with i_{lim} 90
- 6.2 Number of galaxies in the shear catalogue 90
- 6.3 Results from SIS fits to individual dropout catalogues 94
- 6.4 Results from SIS fits to different background galaxy samples 94
- 6.5 New cluster redshifts 100

- 7.1 Number of clusters in the two cluster lists 106
- 7.2 Filter choice wrt redshift range for luminosity calculations 108
- 7.3 Number of galaxies in the photometric redshift catalogue 108

- A.1 Optically selected cluster list 131
- A.2 X-ray selected cluster list 135

List of Tables

Introduction and summary

In 1933 Zwicky noticed that clusters of galaxies contain about 10 times more mass than expected from the luminosities of the visible cluster member galaxies. This was the first strong observational evidence for the existence of a previously unknown (dark) matter component. Comprising the most massive gravitationally bound and relaxed structures in the Universe, galaxy clusters are essential in providing a deeper understanding of the properties of dark matter. Being cosmologically young objects, they are undergoing strong evolution from redshift ~ 1 until today. The cosmological evolution of clusters thus provides a direct insight into the growth of cosmic structures.

Gravitational lensing is an excellent tool for studying the mass distribution in the Universe, because it does not make assumptions about the nature or the dynamical state of the gravitating matter. In particular, lensing does not discriminate between the luminous and dark matter, and is thus a unique tool for studying the latter which otherwise evades observations. This aspect is very important for an analysis of the massive structures in the Universe, since dark matter dominates their state and evolution.

In Chaps. 1 to 4 we give an introduction into cosmology, weak lensing, and clusters of galaxies. We also review data reduction and redshift measurement techniques.

In Chap. 5 we apply weak gravitational lensing to two intermediate-redshift massive galaxy clusters, Abell 1351 and Abell 1995. We investigate their overall mass distribution by deriving two-dimensional mass maps of the clusters, visualising their surface mass distributions. We detect a significant neighbouring mass peak close to the cluster centre of Abell 1351, coinciding with the galaxy distribution and further supported by a faint radio emission filament pointing towards the peak. We also fit predicted shear values from theoretical models to the shapes of the lensed galaxies, and estimate the cluster masses to $M_{200} \sim 8-9 \times 10^{14} h_{70}^{-1} M_{\odot}$ (Abell 1351) and $M_{200} \sim 5-6 \times 10^{14} h_{70}^{-1} M_{\odot}$ (Abell 1995).

Galaxy clusters at high redshifts ($z > 1$) are often still forming and not fully virialised. Applying gravitational lensing to $z > 1$ clusters is thus especially useful in studies of the earlier Universe. Since the abundance and mass of such clusters are extremely sensitive to cosmological parameters, their mass properties should be investigated in detail. In Chap. 6 we study a sample of 11 high-redshift ($1.2 < z < 1.6$) galaxy clusters in the Cosmic Evolution Survey (COSMOS) using weak lensing measurements from deep space-based, high-resolution images. We utilise a very clean high-redshift ($z \gtrsim 2.5$) background sample of Lyman-break galaxies for this purpose, both alone and in combination with galaxies with and without photometric redshift estimates. Due to the low cluster masses

Introduction and summary

we statistically stack the clusters to increase the signal-to-noise of our measurements. The derived mass estimate of $M_{\text{SIS}} = (9.0 \pm 4.7) \times 10^{13} h_{70}^{-1} M_{\odot}$ (within $r = 1.0 h_{70}^{-1}$ Mpc) is nonetheless based on a source catalogue limited by a large contamination from foreground objects (lacking individual redshift information), which is also reflected in the large error bars. Since the photometric redshift catalogue is magnitude limited ($i < 25$), utilising only these galaxies does not yield a significant detection due to the much higher relative lensing efficiency in the non-photometric sample. This is the first time that high redshift ($z > 1$) clusters have been measured with weak lensing based on a sample of Lyman-break background galaxies, and it is one very few studies of high-redshift clusters using weak gravitational lensing in general.

The mass-to-light (M/L) ratio yields the total amount of mass relative to luminosity, thus specifying the relative contribution of the dark matter component. The M/L ratio increases with mass for cosmic objects in the Universe, seemingly approaching a saturation value at cluster scales where it is not increasing with mass anymore. A universal M/L relation inferred from that of galaxy clusters can thus yield constraints on the matter density parameter of the Universe. In Chap. 7 we investigate the M/L ratio of 127 optically selected and 51 X-ray clusters over a wide redshift range ($0.2 \leq z < 1.6$). After correcting the cluster luminosities for passive evolution, we find no significant evolution in the M/L ratio with redshift. Exploring the differences between the two cluster samples we find they generally agree very well. This applies to both redshift independent fits as well as when the datasets have been divided into redshift bins. The results are also in good agreement with the literature.

Chapter 1

Cosmology

Cosmology comprises the scientific study of the Universe as a whole, describing its current state and evolution with time. An essential part of this science encompasses testing theoretical predictions against observations; however, given our unique Universe, cosmology is a distinctive science. There is only one universe, and we are given one set of parameters to investigate. We cannot look at different universes with slightly different initial conditions and correlate their behaviour. While physics is predominantly characterised by experiment and theory, cosmology is mostly characterised by observation and a theoretical framework.

Cosmology as a science did not really mature until the 20th century, when the field experienced tremendous development. Einstein finished his theory of General Relativity in the early 1910's, and in the late 1920's Hubble discovered that nebulae are galaxies of their own and that they are in fact receding with a speed proportional to their distance from the Earth. During the next few decades significant theoretical and observational development took place, including the instigation of the theory of the Big Bang. The cosmic microwave background (CMB) radiation was discovered through radio observations in the mid 1960's – very small fluctuations in the CMB as well as inhabiting a perfect blackbody spectrum were also predicted around this time but not observed until the early 1990's by the COsmic Background Explorer (COBE) satellite. Structure formation became a hot research topic in the 1970's, inflation a decade later.

Having been proposed by Zwicky in the 1930's, dark matter again gained interest in the 1980's with the application of gravitational lensing to galaxy clusters and the idea that dark matter is a new, exotic type of non-baryonic, non-luminous matter. The temperature estimates of clusters through X-ray satellites provided further insight; very high temperatures who could only be explained through a cluster potential well much deeper than expected from visible cluster components alone. Galaxy clusters therefore became the centre of attention with respect to the question of dark matter. In addition they are cosmologically young objects still containing information about their initial conditions, thus they form an important tool with respect to understanding structure formation and the attraction of matter through gravitational instability.

Cosmology forms the framework for clusters of galaxies, the main research topic of this thesis. Here clusters are used as cosmological tools, yielding further insight into the evolution of our Universe, as well as providing important information regarding its

largest matter component – dark matter. In this context, this chapter summarises the theory behind cosmology. Extensive descriptions on the subject can be found in Schneider (2008), Peacock (1999), and Schneider (2006c,a). A comprehensive introduction to General Relativity is given by Hartle (2003).

Note that symbols written in bold denote vectors, i.e. $\mathbf{r} = \vec{r}$, where lowercase letters represent three-dimensional vectors and uppercase letters two-dimensional vectors.

1.1 An expanding Universe

Because of the finite speed of light, $c = 3 \times 10^8 \text{ m s}^{-1}$, observations are limited to sources lying on our backward light cone. In a Euclidean space (i.e. a flat and static Universe) this is represented by the points in spacetime which satisfy $|\mathbf{r}| = r = c(t_0 - t)$, provided that we are located at $\mathbf{r} = 0$ today at $t = t_0$. It is therefore difficult to distinguish between laws of nature and cosmic coincidences. Nevertheless, since the Universe appears to be homogeneous and isotropic on large scales (“the cosmological principle”, see Sect. 1.2), our understanding of it has come a long way, as well as being in constant development.

Hubble (1929b) discovered that the more distant a galaxy is, the faster it appears to be moving away from the observer. He also found a relation between the radial velocity with which galaxies recede from us, v , and their distance, D ,

$$v = H_0 D , \tag{1.1}$$

where H_0 is a constant of proportionality (Hubble 1929a). Equation (1.1) was later called the Hubble law, and the constant named the Hubble constant. We now interpret the Hubble law to predict an expanding Universe, a theory arising from the idea proposed by Hubble himself, that the light-waves from distant objects are affected “by some property of space or by forces acting on the light during its long journey to the Earth” (Hubble 1929b).

Hubble measured v the same way we do today, by taking a spectrum of a galaxy and looking at its spectral lines. If they are shifted towards longer wavelengths (i.e. redshifted) the galaxy is moving away from us, as is the case for most galaxies. This wavelength shift of the spectral lines defines the redshift, z , of an object,

$$z \equiv \frac{\lambda_{\text{obs}} - \lambda_0}{\lambda_0} , \quad \lambda_{\text{obs}} = (1 + z)\lambda_0 . \tag{1.2}$$

Here λ_{obs} is the observed wavelength of a photon emitted by the object, and λ_0 its initial wavelength. Given that the object exhibits strong spectral features it is straightforward to determine an object’s redshift using spectra. Measuring the distance (see Sect. 1.7), however, is the key problem in determining H_0 . When presenting distances based on redshift, they therefore always contain a factor of h^{-1} , where h is the scaled Hubble constant defined as

$$H_0 = h \, 100 \text{ km s}^{-1} \text{ Mpc}^{-1} . \tag{1.3}$$

In the pursuit for H_0 , one of the best results yields (assuming a flat Universe; $\Omega_0 = 1$, see Sect. 1.5)

$$H_0 = (70 \pm 2) \text{ km s}^{-1} \text{ Mpc}^{-1} , \tag{1.4}$$

obtained from the Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) observations (Komatsu et al. 2010). The WMAP satellite mission has been doing full-sky observations of the CMB radiation (Sect. 3.4) since 2001. This value has also been confirmed by other independent studies, e.g. Suyu et al. (2010) or Riess et al. (2009). In the following we present the theoretical framework in which such an expanding Universe can be described.

1.2 Cosmic evolution

Observations reveal that faint galaxies (e.g. $R > 20$; see Sect. 4.3.2 for a definition of filters) are uniformly distributed in the Universe when smoothed on scales $\gtrsim 200$ Mpc, indicating an isotropic distribution of galaxies over large angles. In addition, approximately isotropic CMB radiation is reaching us from all directions, further supporting these observations. By assuming our location in the Universe is not special (generalisation of the Copernican Principle), this large-scale isotropy should be found by all other observers, implying that the Universe is spatially homogeneous and isotropic. The term “the cosmological principle” is used to describe such a universe.

1.2.1 The Universe on small scales

Although General Relativity (see Sect. 1.3) is currently the best description of gravitation, the Newtonian approach is valid on small scales. We will therefore first apply Newtonian gravitation to a small region of a homogeneous and isotropic world model.

Consider a homogeneous sphere of matter; we allow this sphere to radially expand such that the density, $\rho(t)$, is always spatially homogeneous. At a given time t_0 any matter element of the sphere is located at a position \mathbf{x} . At some other time t its location will have changed to $\mathbf{r}(t)$, parallel to the velocity vector of the particle, implying that

$$\mathbf{r}(t) = a(t)\mathbf{x} . \quad (1.5)$$

Because $\rho(t)$ is always homogeneous, $a(t)$ is spatially constant and hence only depends on time. The function $a(t)$ is called the cosmic scale factor, which is normalised such that $a(t_0) = 1$, where t_0 is the current cosmic time. Note that with this normalisation, \mathbf{r} , a , and \mathbf{x} are all dimensionless (which also results in an unusual dimension for the mass of this sphere, see eq. 1.9). The worldline, (\mathbf{r}, t) , for all particles (or observers) is determined by their current position, \mathbf{x} , called the comoving coordinate. The velocity of any such particle is

$$\mathbf{v}(\mathbf{r}, t) = \dot{\mathbf{r}}(t) = \dot{a}(t)\mathbf{x} = \frac{\dot{a}}{a}\mathbf{r} \equiv H(t)\mathbf{r} , \quad (1.6)$$

where the expansion rate is defined as the Hubble parameter,

$$H(t) \equiv \frac{\dot{a}}{a} . \quad (1.7)$$

The present value of $H(t)$ is the Hubble constant, $H_0 \equiv H(t_0)$. This relation can be found from calculating the relative velocity of two particles at \mathbf{r} and $\mathbf{r} + \Delta\mathbf{r}$,

$$\Delta\mathbf{v} = \mathbf{v}(\mathbf{r} + \Delta\mathbf{r}, t) - \mathbf{v}(\mathbf{r}, t) = H(t)\Delta\mathbf{r} , \quad (1.8)$$

which equals eq. (1.1) for $t = t_0$. Equation (1.8) thus generalises the Hubble law to arbitrary time.

1.2.2 The evolution of the cosmic scale factor

To understand how $a(t)$ evolves with time, we look at a spherical surface of radius x at time t_0 , corresponding to $r(t) = a(t)x$ at t . The mass inside this surface, $M(x)$, is constant in time, such that by defining the mass density of the Universe today as $\rho_0 \equiv \rho(t_0)$ we get

$$\begin{aligned} M(x) &= \frac{4\pi}{3}\rho_0 x^3 = \frac{4\pi}{3}\rho(t)r^3(t) \\ &= \frac{4\pi}{3}\rho(t)a^3(t)x^3. \end{aligned} \quad (1.9)$$

Due to mass conservation, the density is inversely proportional to the volume of the sphere,

$$\rho(t) = \rho_0 a^{-3}(t). \quad (1.10)$$

We can derive the equation of motion of a particle on the spherical surface from its gravitational acceleration towards the centre of the sphere as

$$\ddot{r}(t) = -\frac{GM(x)}{r^2} = -\frac{4\pi G}{3} \frac{\rho_0 x^3}{r^2}, \quad (1.11)$$

where $G = 6.7 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$ is the gravitational constant. Substituting for $r(t) = xa(t)$ we get

$$\ddot{a}(t) = -\frac{4\pi G}{3} \frac{\rho_0}{a^2(t)} = -\frac{4\pi G}{3} \rho(t)a(t). \quad (1.12)$$

Multiplying eq. (1.12) with $2\dot{a}$ and integrating with respect to time yields the (first) Friedmann equation,

$$\dot{a}^2 = \frac{8\pi G}{3} \rho_0 \frac{1}{a} - Kc^2 = \frac{8\pi G}{3} \rho(t)a^2(t) - Kc^2, \quad (1.13)$$

where Kc^2 is the constant of integration and K is called the curvature parameter (see below).

We can multiply eq. (1.13) with $mx^2/2$ to see that this equation describes the conservation of energy, namely

$$\frac{mv^2(t)}{2} - \frac{GMm}{r(t)} = -Kmc^2 \frac{x^2}{2}. \quad (1.14)$$

The first term represents the kinetic energy, E_{kin} , of a particle of mass m on the spherical surface. The potential energy, E_{pot} , of the particle is given by the gravitational potential of the sphere. From eq. (1.14) we therefore know that the total energy of the particle, $E_{\text{tot}} = E_{\text{kin}} + E_{\text{pot}}$, is constant. It is important to note that K is proportional to E_{tot} , meaning that the history of the expansion depends on K , in particular on the sign of K :

$K > 0$ Closed universe: the Universe will expand until $a = a_{\text{max}} = 8\pi G\rho_0/(3Kc^2)$ (where $\dot{a} = 0$), after which it will re-collapse as the expansion turns into a contraction.

$K = 0$ Flat universe: the Universe will expand forever, but such that $\dot{a} \rightarrow 0$ for $t \rightarrow \infty$.

$K < 0$ Open universe: the Universe will expand forever.

Note that the considerations given in this section apply to a matter-dominated Universe only ($\Omega_\Lambda = \Omega_r = 0$, see also Sect. 1.5).

1.3 General Relativity

In 1915, Einstein formulated his theory of gravitation, General Relativity (GR), which is considered to be the most accurate theory explaining the physics of gravitation¹ (see also Sect. 2.1.1). In 1916 he published his field equations, describing gravitation as a result of spacetime being curved by matter and energy (Einstein 1916). Like most of his contemporaries in the early 20th century, Einstein believed the Universe to be static. Since his field equations did not allow for a solution corresponding to a homogeneous, isotropic, and static universe, Einstein later generalised his equations, including an additional term, the cosmological constant, Λ . With the discovery that the Universe is expanding, however, Einstein subsequently discarded this additional term. Nevertheless, observations from the last decade show that $\Lambda \neq 0$ (see Sect. 1.8.5).

Most of the equations in Sect. 1.2 are still valid within the framework of GR, although their interpretation has changed considerably. Note that neither the Friedmann equation (1.13), the equation of motion (1.12), or the evolution of the matter density (1.10) make any reference to a sphere. In a homogeneous and isotropic universe all points are equal. Unlike an expanding sphere, the Universe is not contained within a boundary, nor does it have a centre or any single point in the current space-like hypersurface from which everything is moving away. Specifically, in GR the expansion is interpreted in the sense that the space itself is expanding, carrying particles with it. Furthermore, such an expansion has implications for photons travelling in spacetime; as time advances and space expands, the wavelength of a photon is stretched. In other words, the younger the Universe was when the photon was emitted, the farther it has travelled, and the more redshifted it will be. It is important to note that this redshifting is not the same as Doppler redshift, but rather a property of the expanding Universe.

1.3.1 Obtaining the Friedmann equation from General Relativity

GR describes gravitation as a curved spacetime geometry rather than a force. Einstein's field equations,

$$G_{\alpha\beta} + \Lambda g_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta} , \quad (1.15)$$

relate the spacetime curvature (expressed by the Einstein tensor, $G_{\alpha\beta}$) to the density and flux of energy and momentum in spacetime (described by the stress-energy tensor, $T_{\alpha\beta}$), concluding that $T_{\alpha\beta}$ is the source of curvature (i.e. the generalisation of the stress tensor

¹There are other theories in which Newtonian physics are modified in order to reduce the amount of dark matter needed, such as MODified Newtonian Dynamics (MOND). Since MOND is not able to accurately explain the dynamics of massive structures in the Universe, e.g. galaxy clusters which is the main topic of this thesis, without also introducing dark matter, we will not explore this theory further.

in Newtonian physics to GR). In eq. (1.15) G is the gravitational constant and c denotes the speed of light.

The field equations (1.15) must be solved in order to obtain the metric tensor $g_{\alpha\beta}$. However, since $G_{\alpha\beta}$ is a complicated function of $g_{\alpha\beta}$, this is not a trivial task. Only a handful of analytic solutions are known, requiring highly symmetric mass distributions. By re-writing the metric $g_{\alpha\beta}$ as a line element,

$$ds^2 = \sum_{\alpha,\beta=0}^3 g_{\alpha\beta} dx^\alpha dx^\beta, \quad (1.16)$$

Robertson and Walker found that what is now called the Robertson-Walker metric,

$$ds^2 = c^2 dt^2 - a^2(t) [dx^2 + f_K^2(x) (d\theta^2 + \sin^2 \theta d\varphi^2)], \quad (1.17)$$

solves the field equations for a spatially homogeneous and isotropic world model. Here, t is the cosmic time (equal to the time measured by comoving observers, see Sect. 1.6), x is the comoving coordinate, and θ and φ are the angular coordinates on a unit sphere. The comoving angular diameter distance, $f_K(x)$ (see Sect. 1.7), depends on the curvature parameter, K , as

$$f_K(x) = \begin{cases} K^{-1/2} \sin(K^{1/2}x) & (K > 0) \\ x & (K = 0) \\ (-K)^{-1/2} \sinh[(-K)^{1/2}x] & (K < 0) \end{cases}. \quad (1.18)$$

In other words, for a three-dimensional space of constant curvature K , its spherical coordinates are given by (x, θ, φ) .

In a Universe obeying the cosmological principle, the stress-energy tensor describes the matter and energy as $\mathcal{T} = \text{diag}(\rho, p, p, p)$, where $\rho = \rho(t)$ denotes the homogeneous density and $p = p(t)$ the pressure. Inserting the Robertson-Walker metric (1.17) into the field equations (1.15) then yields the (first) Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda c^2}{3}. \quad (1.19)$$

World models for which the Robertson-Walker metric (1.17) applies, and where the cosmic scale factor obeys the Friedmann equation (1.19), are called Friedmann-Lemaître (FL) models.

1.3.2 Obtaining the equation of motion from General Relativity

The equation of motion is inferred by multiplying with a^2 and differentiating eq. (1.19) with respect to time, yielding

$$2\dot{a}\ddot{a} = \frac{8\pi G}{3}(\dot{\rho}a^2 + 2a\dot{a}\rho) + 2a\dot{a}\frac{\Lambda c^2}{3}. \quad (1.20)$$

Re-writing the first law of thermodynamics (eq. (1.24)) as

$$\dot{\rho}a^3 + 3\rho a^2\dot{a} = -\frac{p}{c^2}3a^2\dot{a} \implies \dot{\rho}a^2 = -3a\dot{a}\left(\rho + \frac{p}{c^2}\right), \quad (1.21)$$

we can eliminate $\dot{\rho}$ from eq. (1.20), resulting in the equation of motion

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2} \right) + \frac{\Lambda c^2}{3}. \quad (1.22)$$

Equation (1.22) is together with eq. (1.19) called the Friedmann equations.

Comparing eq. (1.22) to the Newtonian equation of motion (1.12), we see that the pressure term is missing from the latter equation. In addition, a term including the cosmological constant is added to eq. (1.22). This follows directly from GR and Einstein's generalisation of his field equations to allow for a static world model (see the beginning of Sect. 1.3).

1.4 Matter models of the Universe

The metric (1.17) describes a homogeneous perfect fluid with density $\rho(t)$ and pressure $p(t)$. An individual galaxy is then thought of as a particle in this fluid. Our homogeneous and isotropic Universe consists of three non-interacting² matter components: pressureless matter, radiation, and vacuum energy. The term pressureless matter here encompasses matter where its components have random velocities $\ll c$, for which $p \ll \rho c^2$. Such pressure is insignificant in the cosmological context, and we can approximate $p_m = 0$. With random motions yielding a thermal energy much less than the rest energy, the matter in the Universe can be described as pressureless. By contrast, radiation is characterised by massless photons moving in random directions with the speed of light, implying that $p = \rho c^2/3$ from the equation of state for radiation.

The evolution of the density, ρ , depends on the equation of state of the matter, characterised by a dimensionless number equal to the ratio of its pressure, p , to its energy density. We know from special relativity that the energy density of ordinary matter is ρc^2 . The energy within a sphere of radius r is then $U = 4\pi\rho c^2 r^3/3$. The first law of thermodynamics states that the change in volume, dV (where $V = 4\pi r^3/3$), causes a change in internal energy, dU , equal to the work $dU = -pdV$. Applying the equations of GR (see Sect. 1.3) to our homogeneous and isotropic Universe, this relation becomes

$$d(\rho c^2 r^3) = -p d(r^3). \quad (1.23)$$

Inserting for eq. (1.5) we get

$$d(\rho c^2 a^3) = -p d(a^3). \quad (1.24)$$

Applying eq. (1.24) to pressureless matter, $p_m = 0$, yields

$$\rho_m(t) = \rho_{m,0} a^{-3}(t), \quad (1.25)$$

which was also found from the conservation of mass in eq. (1.10). Recall that $\rho_{i,0} = \rho_i(t_0)$, where $i \in \{m, r\}$. Inserting $p_r = \rho_r c^2/3$ into eq. (1.24) renders the corresponding equation for radiation,

$$\frac{d(\rho_r a^3)}{da} = -\frac{\rho_r}{3} \frac{d(a^3)}{da}. \quad (1.26)$$

²Although these matter components did interact in the very early Universe, we assume they have been independent during most of its history.

Chapter 1. Cosmology

Assuming the solution is of the form (1.25), we replace ρ_r with $\rho_{r,0}a^n$ in eq. (1.26), leading to $(3+n)a^{2+n} = -a^{2+n} \Rightarrow n = -4$, which yields

$$\rho_r(t) = \rho_{r,0}a^{-4}(t) . \quad (1.27)$$

From this result we know that the energy density of radiation scales with a^{-4} . This is in part due to the decrease $\propto a^{-3}$ in number density of the radiation elements (e.g. photons) like for pressureless matter. A second contribution comes from the redshifting of photons. The energy of a photon is $E = h_P\nu = h_Pc/\lambda$, where h_P is the Planck constant, ν the frequency of the photon, and λ its wavelength. Since the wavelength of a photon changes proportionally to a (eq. (1.44)), its energy changes as $\propto a^{-1}$, resulting in a total decrease in photon energy density $\propto a^{-4}$.

Finally, imagine that empty space has a finite energy density called the vacuum energy density, ρ_Λ . One would then assume that ρ_Λ remains constant in space and time, $d\rho_\Lambda/dt = 0$. Inserting this into eq. (1.24) gives

$$p_\Lambda = -\rho_\Lambda c^2 , \quad (1.28)$$

in other words, vacuum energy density corresponds to a negative pressure.

The sum of the three matter density (pressure) components then yields the total density (pressure) of the Universe,

$$\rho = \rho_m + \rho_r + \rho_\Lambda = \rho_{m,0}a^{-3} + \rho_{r,0}a^{-4} + \rho_\Lambda \quad (1.29)$$

$$p = \frac{\rho_r c^2}{3} - \rho_\Lambda c^2 = \frac{\rho_{r,0}c^2}{3a^4} - \rho_\Lambda c^2 . \quad (1.30)$$

1.5 Cosmological parameters

The critical density, ρ_{cr} , is defined as the density today's Universe would have if it was flat ($K = 0$). It is derived from eq. (1.19) by letting $t = t_0$ and $H_0 = \dot{a}(t_0)$ (since $a(t_0) = 1$), and defining ρ according to eq. (1.29) (i.e. without a separate Λ term),

$$\rho_{cr} \equiv \frac{3H_0^2}{8\pi G} \approx 9.22 \times 10^{-30} h_{70}^2 \text{ g cm}^{-3} . \quad (1.31)$$

The density parameters are defined in terms of ρ_{cr} as

$$\Omega_m \equiv \frac{\rho_{m,0}}{\rho_{cr}} , \quad \Omega_r \equiv \frac{\rho_{r,0}}{\rho_{cr}} , \quad \Omega_\Lambda \equiv \frac{\rho_\Lambda}{\rho_{cr}} = \frac{\Lambda c^2}{3H_0^2} . \quad (1.32)$$

The energy density of electromagnetic radiation in the Universe today is dominated by the CMB energy density. Since $\Omega_{\text{CMB}} \simeq 4.9 \times 10^{-5} h_{70}^{-2}$ (calculated from the Stefan-Boltzmann law), the effect of radiation on the expansion rate can be neglected for the current epoch. However, since the radiation density drops as a^{-4} and the matter density as a^{-3} , there must have been a time, a_{eq} , at which matter and radiation had the same energy density. The Universe was therefore radiation dominated for $a \lesssim a_{\text{eq}}$.

Using the Friedmann equation (1.19) and inserting for eqs. (1.29) and (1.32) render

$$H^2 = H_0^2 \left[\frac{\Omega_r}{a^4} + \frac{\Omega_m}{a^3} + \frac{Kc^2}{a^2 H_0^2} + \Omega_\Lambda \right] . \quad (1.33)$$

Applying eq. (1.33) to the current epoch, $a = 1$, provides an expression for the curvature of our present-day Universe, $K = (\Omega_m + \Omega_\Lambda - 1)H_0^2/c^2$ (using $\Omega_r \ll \Omega_m$). Defining the total density parameter of today as $\Omega_0 = \Omega_m + \Omega_\Lambda + \Omega_r$ and inserting into eq. (1.33) now yields the expansion equation

$$H^2 = H_0^2 [\Omega_r a^{-4} + \Omega_m a^{-3} + (1 - \Omega_0) a^{-2} + \Omega_\Lambda] . \quad (1.34)$$

We see from the above expressions that the curvature of the Universe is specified by the total density, i.e.

$$\begin{aligned} \Omega_0 > 1 &\iff K > 0 , \\ \Omega_0 = 1 &\iff K = 0 , \\ \Omega_0 < 1 &\iff K < 0 . \end{aligned}$$

Since $H(t) = \dot{a}(t)/a(t)$, eq. (1.34) can be integrated using $a(t_0) = 1$ to obtain the evolution of $a(t)$ with time. Models including a minimum $0 < a = a_{\min} < 1$, at which the Universe went from contracting ($H < 0$) to expanding ($H > 0$), predicts a maximum redshift, $z_{\max} = 1/a_{\min} - 1$ (see eq. (1.44)). From observations we know that sources at redshifts as large as 8.6 exist (Lehnert et al. 2010), leading to the condition $a_{\min} < 0.10$, which again requires Ω_m to be very small. As observations of galaxies and galaxy clusters have determined $\Omega_m > 0.1$, such an $a_{\min} > 0$ cannot exist. Thus, for $a < 1$, the cosmic scale factor is monotonically increasing with time, and the Universe must have experienced a time in the past at which $a = 0$ (so-called Big Bang models). Since both matter and radiation density diverge as $a \rightarrow 0$, the density at this point is singular in this formalism.

1.5.1 Future expansion of the Universe

We know from the redshift of galaxies that the Universe is expanding, i.e. $H_0 > 0$. Utilising eq. (1.34) we can predict how the cosmic expansion further evolves with time, for $a > 1$. The values of a for which H changes sign evidently coincides with a vanishing right-hand side of the equation. Since $\Omega_r < 10^{-4}$ we can ignore radiation, and eq. (1.34) can then be written as

$$f(a) \equiv \frac{H^2(a)}{H_0^2} a^3 = \Omega_m + (1 - \Omega_m - \Omega_\Lambda)a + \Omega_\Lambda a^3 . \quad (1.35)$$

We will first look at solutions for $\Omega_\Lambda = 0$. Equation (1.35) then becomes $f(a) = \Omega_m + (1 - \Omega_m)a$, for which a solution exists for all $\Omega_m \neq 1$. However, as $\Omega_m < 1$ requires $a < 0$, which is unphysical, solutions only make sense for $\Omega_m > 1$ and we find that $f(a)$ vanishes for

$$a_{\text{coll}} = \frac{\Omega_m}{\Omega_m - 1} > 1 . \quad (1.36)$$

This means that for $\Omega_\Lambda = 0$, $\Omega_m > 1$, the Universe will reach a maximum expansion at $a = a_{\text{coll}} > 1$, after which it will re-collapse.

The second case covers models for a flat Universe, for which $\Omega_m + \Omega_\Lambda = 1$. Equation (1.35) then becomes $f(a) = \Omega_m + (1 - \Omega_m)a^3$. In the case where $\Omega_m \leq 1$ (and hence

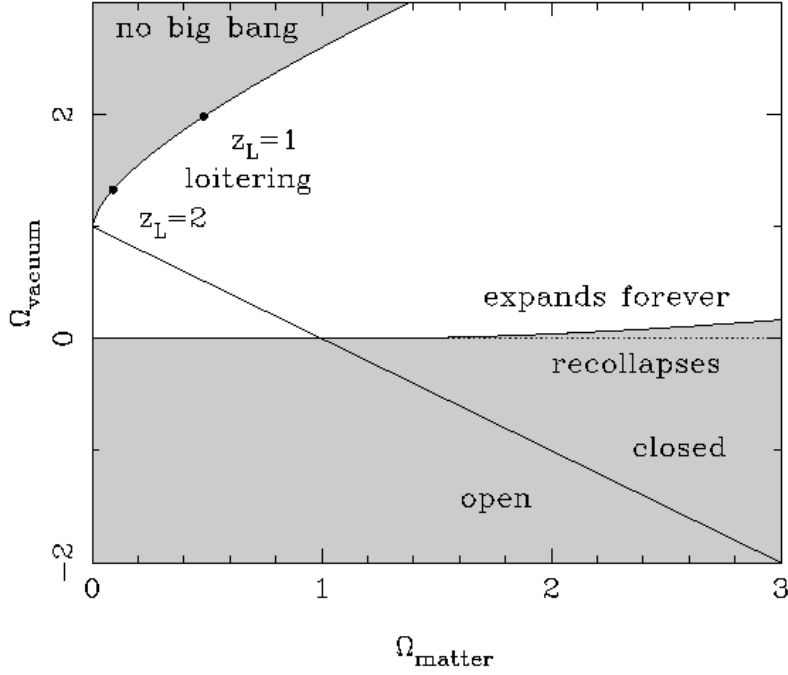


Figure 1.1: Classification of cosmological models as a function of Ω_Λ (denoted Ω_{vacuum} in the plot) and Ω_m (denoted Ω_{matter}). The straight diagonal line running from the upper left to the lower right shows the solution for a flat universe ($\Omega_m + \Omega_\Lambda = 1$), below and above which the open ($K < 0$) and closed ($K > 0$) universe solutions lie, respectively. The shaded area in the lower part of the figure represents the models for which the Universe will re-collapse, for the models lying above this area the Universe will expand forever. The shaded area in the upper left corner represents the models predicting a maximum redshift and hence no Big Bang. We know from observations that these models can be excluded, see the text for more details. (Figure credit: Peacock 1999)

$\Omega_\Lambda \geq 0$), $f(a) = 0$ has no solution, and the Universe will expand forever. For $\Omega_m > 1$ a solution is found for

$$a_{\text{coll}} = \left(\frac{\Omega_m}{\Omega_m - 1} \right)^{1/3} > 1, \quad (1.37)$$

in which the Universe will re-collapse after $a = a_{\text{coll}}$. Figure 1.1 describes possible expansion histories as a function of Ω_m and Ω_Λ .

By defining $a = 0$ as the beginning of time, $t = 0$, the age of the Universe is defined as the time since the Big Bang. Since $a(t) < 1$ is monotonically increasing, the scale factor can be used to describe cosmic time. Ignoring Ω_r and applying

$$dt = \frac{da}{\dot{a}} = \frac{da}{aH} \quad (1.38)$$

to eq. (1.34), we find that

$$t(a) = \frac{1}{H_0} \int_0^a da' [\Omega_m a'^{-1} + (1 - \Omega_m - \Omega_\Lambda) + \Omega_\Lambda a'^2]^{-1/2}. \quad (1.39)$$

The current age of the Universe, $t = t_0$, can then be calculated by solving for $a = 1$, yielding $t_0 \sim H_0^{-1}$ (ignoring a numerical factor which depends on the density parameters).

1.6 Redshift in the context of cosmic expansion

Massless particles like photons propagate along null geodesics. In GR's curved spacetime, a geodesic means the same as a straight line does in Newtonian physics. In other words, a particle free from external forces will have a worldline equal to a geodesic, where a worldline is the path of a particle moving through spacetime. A null geodesic is defined as having $ds^2 = 0$ in the metric (1.17). In addition, radial curves, for which $\theta = \text{const.}$, $\varphi = \text{const.}$, are spatial projections of geodesics. Radial light rays therefore have $c^2 dt^2 = a^2 dx^2$, as seen from eq. (1.17). Furthermore, by choosing our location to be at the centre of the coordinate system, $x = 0$, we can write

$$c dt = -a(t) dx . \quad (1.40)$$

The minus sign comes from the fact that the photons reaching us have $dt > 0$ but $dx < 0$. Their radial coordinates can be calculated by integrating over cosmic time, from eq. (1.5), as

$$x(t) = \int_t^{t_0} \frac{c dt'}{a(t')} . \quad (1.41)$$

Any source observed today hence has a radial coordinate given by eq. (1.41).

A comoving source at distance x from us emits radiation at t and $t + \Delta t_e$, reaching us at t_0 and $t_0 + \Delta t_{\text{obs}}$. The comoving distance will remain constant, yielding

$$\int_t^{t_0} \frac{c dt'}{a(t')} = x = \int_{t+\Delta t_e}^{t_0+\Delta t_{\text{obs}}} \frac{c dt'}{a(t')} = \int_t^{t_0} \frac{c dt'}{a(t')} + \frac{c \Delta t_{\text{obs}}}{1} - \frac{c \Delta t_e}{a(t)} , \quad (1.42)$$

where, in the last step, we assume that Δt is small and $a(t) \approx a(t + \Delta t)$, along with inserting for $a(t_0) = 1$. From eq. (1.42) we see that

$$\Delta t_e = a(t) \Delta t_{\text{obs}} . \quad (1.43)$$

Consequently, two instances occurring in the rest-frame of the source are measured with a time separation of Δt_e , but will arrive at the observer's rest-frame having a separation of $\Delta t_{\text{obs}} = \Delta t_e / a(t) > \Delta t_e$ given that the Universe is expanding. Further, since the frequency of the emitted radiation, ν , is the inverse of Δt_e , the corresponding relation also holds for the observed radiation, thus $\nu_e = \nu_{\text{obs}} / a(t) > \nu_{\text{obs}}$. Using eq. (1.2) we can then write

$$1 + z \equiv \frac{\lambda_{\text{obs}}}{\lambda_e} = \frac{\nu_e}{\nu_{\text{obs}}} = \frac{1}{a(t)} . \quad (1.44)$$

All photons observed from comoving sources are therefore redshifted.

Section 4.4 describes how to measure the redshift of an object using spectroscopic and photometric observations.

1.7 Distance measurements in a dynamic spacetime

As our (possibly) curved Universe expands and time passes, distances in this spacetime change. In observations we look along our backward light cone, as opposed to distances

measured at a given instant of time. The measure “distance” therefore does not have one unique meaning. All of a , t , and z can be used to describe distance. However, since z is the only parameter that can actually be measured, extragalactic distances are usually expressed in terms of redshift.

Given a source at redshift z with angular diameter δ and physical diameter $2R$, its angular diameter distance is given as the ratio between the two,

$$D_{\text{ang}}(z) = \frac{2R}{\delta} = a(z)f_K(x). \quad (1.45)$$

From eq. (1.40) we see that

$$-dx = \frac{c dt}{a} = \frac{c da}{a \dot{a}} = \frac{c da}{a^2 H}, \quad (1.46)$$

which, inserted into the metric (1.17) together with setting $ds = 2R$ and $d\theta = \delta$, yields the last equality in eq. (1.45). We can also use eq. (1.46) to derive an expression for the comoving distance between two sources as a function of redshift

$$\begin{aligned} x(z_1, z_2) &= \frac{c}{H_0} \int_{a(z_2)}^{a(z_1)} [a\Omega_m + a^2(1 - \Omega_m - \Omega_\Lambda) + a^4\Omega_\Lambda]^{-1/2} da \\ &= x(z_2) - x(z_1), \end{aligned} \quad (1.47)$$

where $z_1 < z_2$ and $x(z) \equiv x(0, z)$ (the comoving distance between us and a source at z). Equation (1.45) can be generalised to yield the angular diameter distance between two sources as seen from redshift $z_1 < z_2$,

$$D_{\text{ang}}(z_1, z_2) = a(z_2)f_K[x(z_1, z_2)]. \quad (1.48)$$

Note that $D_{\text{ang}}(z_1, z_2) \neq D_{\text{ang}}(z_2) - D_{\text{ang}}(z_1)$.

In addition to the one described above, a second distance measurement method exists. Here the observed flux, S , of a source is related to its luminosity, L , rendering the luminosity distance to the source,

$$D_{\text{lum}}(z) \equiv \sqrt{\frac{L}{4\pi S}}. \quad (1.49)$$

We see that $D_{\text{lum}} = D_{\text{ang}}$ in a static Universe, whereas in a dynamic Universe this only holds for $z \ll 1$ where the expansion is negligible. In general, we can relate the two distance measurements using

$$D_{\text{lum}}(z) = (1+z)^2 D_{\text{ang}}(z) = (1+z)f_K(x). \quad (1.50)$$

In order for eq. (1.50) to be correct, S and L has to be integrated over all frequencies (bolometric). This is because the specific frequency we observe, ν_{obs} , has been redshifted with respect to the emitted frequency, $\nu_e = (1+z)\nu$. The luminosity of the source relates directly to the latter, hence a frequency shift must be taken into account when calculating L from a specific flux. The corrections applied in Sect. 7.3.1, includes this when calculating the luminosity of galaxies.

Finally, the proper distance describes the distance between two comoving objects at a specific instant in time (as opposed to comoving distance, which factors out the expansion of the Universe and hence doesn't change over time). It is defined locally, where space is approximately Euclidean, as seen by an observer close to the sources. We can calculate the proper distance between these two comoving sources from their redshifts, z and $z + \Delta z$ (where $\Delta z \ll 1$), and their angular separation on the sky, $\Delta\theta$ (must be small). The proper separation transverse to the line-of-sight is $D_{\text{ang}}(z)\Delta\theta$. The separation along the line-of-sight is

$$\begin{aligned} \Delta r_{\text{prop}} &= a(z)\Delta x = a(z)\frac{dx}{da}\frac{da}{dz}\Delta z = \frac{ca(z)}{H(z)}\Delta z \\ &= \frac{c}{H_0} \frac{a \Delta z}{\sqrt{\Omega_m a^{-3} + (1 - \Omega_m - \Omega_\Lambda)a^{-2} + \Omega_\Lambda}}, \end{aligned} \quad (1.51)$$

where we have used eqs. (1.46), (1.44), and (1.34).

1.8 The concordance model, Λ CDM

The concordance model is a specific FL model which attempts to explain the existence and structure of the CMB radiation, the accelerating expansion of the Universe, the large-scale structure outlined by galaxies and clusters of galaxies, and the distribution of hydrogen, helium, and lithium in the Universe. It is generally called the Lambda Cold Dark Matter (Λ CDM) model, but is also referred to as the standard cosmological model. The concordance model is in general agreement with the observed Universe, and is the simplest model as such.

The concordance model states that only $\sim 4\%$ of the total density in the Universe today is ordinary (visible) matter. The rest is attributed to dark matter ($\sim 22\%$) and dark energy ($\sim 74\%$). The existence of dark energy is due to a non-vanishing cosmological constant in this model, as described throughout this chapter. Dark matter is introduced in the following section.

Along with all FL models, the concordance model contains two problems; the flatness problem and the horizon problem (Sects. 1.8.2 and 1.8.3). Inflation is presented as a solution to these problems, in Sect. 1.8.4.

1.8.1 Evidence of dark matter

Observing the Coma cluster in 1933, Zwicky discovered a discrepancy between its mass and velocity dispersion (Zwicky 1933, 1937). Assuming the mass-to-light ratio of the Sun, M_\odot/L_\odot , could be used as a proxy for the M/L of all stars, he calculated the mass of the cluster from the luminosities of its galaxies. From the cluster mass he could then calculate the escape velocity of the cluster, i.e. the velocity a galaxy needs to escape from the cluster's gravitational potential. Measuring the radial velocities of the cluster galaxies, he found that their individual velocities were much larger than the escape velocity. Zwicky's conclusion was that the sum of the cluster galaxy masses only accounts for $\sim 10\%$ of the total mass the Coma cluster must have to maintain its high velocity dispersion. There must furthermore be some form of additional matter embedded in the cluster, not emitting

light, but fully dominating the gravitational field in the cluster. This important discovery was the first indicator of the existence of what we now call dark matter.

Since the beginning of X-ray satellites, Zwicky's hypothesis about dark matter has been confirmed by the discovery of hot ($10^7 - 10^8$ K) X-ray gas in clusters. The temperature of the gas renders another measure for the total mass of the clusters, as the hot X-ray gas requires a deep potential well to prevent it from evaporating (see Sect. 3.3). The masses inferred are in agreement with those from velocity dispersion measurements. Note that the mass fraction of the hot X-ray gas is larger than that of the stellar mass only, however it is still much smaller than the total mass required.

The density and temperature of X-ray gas in galaxy clusters can also be measured through the so-called Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zel'dovich 1972). CMB photons are scattered to higher energies while crossing the hot X-ray halo of a cluster, causing it to cast a shadow onto the CMB at low frequencies. The intensity of the SZ effect summed over an entire cluster yields a proxy for the total mass of the cluster.

Gravitational lensing provides a third independent method for determining cluster masses, see Chap. 2. It describes how light is deflected in a gravitational field. By measuring how the light from background galaxies is deflected when passing through a galaxy cluster, the mass of the cluster can be inferred, since the level of deflection depends on how massive the cluster is. The masses derived from lensing effects in clusters of galaxies are generally in agreement with those of the methods described above. However, differences do exist as both X-ray and SZ analyses have to make assumptions about e.g. spherical shape or dynamic relaxations, which do not necessarily hold. We will investigate applications of the gravitational lensing effect in clusters in greater detail in Chaps. 5, 6, and 7.

1.8.2 The flatness problem

At early times, $a \ll 1$, the evolution of the matter density parameter is given by

$$\Omega_m(a) = \frac{\rho_m(a)}{\rho_{cr}(a)} = \frac{8\pi G \rho_m(a)}{3H^2(a)} = \frac{8\pi G}{3H^2(a)} \rho_{m,0}(a) a^{-3} = \left(\frac{H_0}{H}\right)^2 \Omega_{m,0} a^{-3}, \quad (1.52)$$

where we have used generalised versions of eqs. (1.32) and (1.31), together with eq. (1.25) and eq. (1.31) applied for $t = t_0$. Similarly, the evolution of the total density parameter yields

$$\begin{aligned} \Omega_{tot}(a) &= \Omega_r(a) + \Omega_m(a) + \Omega_\Lambda(a) \\ &= \left(\frac{H_0}{H}\right)^2 (\Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_\Lambda) \\ &= \frac{\Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_\Lambda}{\Omega_{r,0} a^{-4} + \Omega_{m,0} a^{-3} + \Omega_\Lambda - (\Omega_{tot,0} - 1) a^{-2}} \\ &= 1 + \frac{\Omega_{tot,0} - 1}{1 - \Omega_{tot,0} + \Omega_{r,0} a^{-2} + \Omega_{m,0} a^{-1} + \Omega_\Lambda a^2} \\ \implies (\Omega_{tot} - 1)(a) &= \frac{(\Omega_{tot} - 1)_0}{1 - \Omega_{tot,0} + \Omega_{r,0} a^{-2} + \Omega_{m,0} a^{-1} + \Omega_\Lambda a^2}, \end{aligned} \quad (1.53)$$

where we have inserted for $(H/H_0)^2$ from eq. (1.33).

Equation (1.53) introduces some interesting questions. First of all we see that $\Omega_{\text{tot}} \rightarrow 1$ for $a \rightarrow 0$, which means that, regardless the value of $\Omega_{\text{tot},0}$, at early times all universes are flat. In addition, except for a minimum value of $a = a_{\text{min}}$, the denominator is larger than zero. In Sect. 1.5 we excluded the possibility of an a_{min} , meaning that an open/closed universe remains open/closed.

Furthermore, applying eq. (1.53) to the very early Universe, $a \ll 1$, we get

$$\Omega(a) - 1 = a^2 \frac{(\Omega_0 - 1)}{\Omega_{r,0}}, \quad (1.54)$$

implying that $|\Omega(a) - 1| \ll 1$ at early times. In order for $(\Omega_0 - 1) \in [-0.9, 2]$, $\Omega(a)$ must have been very close to unity for $a \ll 1$, and must therefore have been subject to a very precise fine-tuning.

1.8.3 The horizon problem

Another problem arises from the fact that different parts of the Universe have very similar physical properties. Given their large separation and the finite speed of light, they can never have been in contact or exchanged information. The problem hence arises as to e.g. why the temperature of the CMB radiation shows such high level of isotropy.

The finite speed of light also determines the size of the visible Universe. Our observable Universe consists of the area from which light has had time to reach us, $t < t_0 \approx 13.5$ Gyr, hence it has a (light travel time) radius of approximately 13.5×10^9 light years. Everything beyond this horizon is not visible to us today.

The finite distance light can travel during the time interval dt , is $c dt$. From eq. (1.5) we see that the corresponding comoving distance interval is $dx = c dt/a$. The comoving distance that light has travelled from the Big Bang to a time t is then given by

$$r_{\text{H,com}}(z) = x = \int_0^t \frac{c dt}{a} = \int_0^a \frac{c da}{a^2 H(a)}. \quad (1.55)$$

For $a_{\text{eq}} \ll a \ll 1$ the expansion rate is dominated by matter, $H(a) \approx H_0 \sqrt{\Omega_m} a^{-3/2}$ (from eq. (1.34)), which is also the main contribution to the integral above. Equation (1.55) can then be written

$$r_{\text{H,com}}(z) \approx \frac{2c}{H_0 \sqrt{\Omega_m}} \sqrt{a} \quad \text{for } a_{\text{eq}} \ll a \ll 1. \quad (1.56)$$

For $a \ll a_{\text{eq}}$ H is radiation-dominated, $H(a) \approx H_0 \sqrt{\Omega_r} a^{-2}$, leading to

$$r_{\text{H,com}}(z) \approx \frac{c}{H_0 \sqrt{\Omega_r}} a \quad \text{for } a \ll a_{\text{eq}}, \quad (1.57)$$

and as expected we see that $r_{\text{H,com}}$ becomes smaller with a .

The recombination epoch (also called the time of last scattering), $z_{\text{rec}} \sim 1000$, was the time when the CMB photons stopped interacting with matter and became free to move around. Afterwards, for $a_{\text{eq}} \ll a \ll 1$, the Universe was dominated by matter. The proper horizon size (see eq. (1.51)) corresponding to eq. (1.56) is then given as

$$r_{\text{H,prop}}(z_{\text{rec}}) = a r_{\text{H,com}}(z_{\text{rec}}) = \frac{2c}{H_0 \sqrt{\Omega_m}} a^{3/2}. \quad (1.58)$$

Chapter 1. Cosmology

In a matter-dominated universe $\Lambda = 0$ applies, for which we can use the Mattig (1958) relation which provides a connection between D_{ang} and z ,

$$D_{\text{ang}}(z) = \frac{2c [\Omega_m z + (\Omega_m - 2) (\sqrt{1 + \Omega_m z} - 1)]}{H_0 \Omega_m^2 (1 + z)^2}. \quad (1.59)$$

Using eq. (1.59), the angular diameter distance to the last scattering surface of the CMB reads

$$D_{\text{ang}}(z) \approx \frac{2c}{H_0 \Omega_m z} \quad \text{for } z \gg 1. \quad (1.60)$$

This length corresponds to an angular size on the sky,

$$\theta_{\text{H,rec}} = \frac{r_{\text{H,prop}}(z_{\text{rec}})}{D_{\text{ang}}(z_{\text{rec}})} \approx \sqrt{\frac{\Omega_m}{z_{\text{rec}}}} \sim \frac{\sqrt{\Omega_m}}{30} \sim \sqrt{\Omega_m} 2^\circ \quad \text{for } \Omega_\Lambda = 0. \quad (1.61)$$

The CMB radiation shows relative fluctuations of $\Delta T/T \sim 10^{-5}$, and is thus very isotropic. From eq. (1.61) we see that regions separated by more than 2 degrees have never been in casual contact before recombination. The horizon problem rises the question of how these regions can have the same temperature, when they have never had the chance to exchange information.

1.8.4 Inflation

The expansion of the Universe can be traced back to a singular state using Einstein's field equations of GR. However, GR does not require the Universe to be homogeneous and isotropic. Inflation was therefore introduced; a theory able to solve the problems that GR could not account for.

The temperature of the Universe increases as we move backwards in time. Approaching the Big Bang, $T \rightarrow \infty$ as $t \rightarrow 0$, reaching energy levels $\sim 10^{14}$ GeV characterised by the unification of the fundamental forces other than gravitation (electromagnetic force, weak and strong interactions). Hence at these energy scales, occurring at $t \sim 10^{-34}$ s, we expect the occurrence of new phenomena, from parts of particle physics which have not yet been fully developed.

Inflation is characterised as the time period in the early Universe where the expansion behaved exponentially, $\ddot{a} > 0$. During this time the expansion was dominated by vacuum energy density, such that

$$a(t) \propto \exp(H_{\text{infl}} t) = \exp\left(\sqrt{\frac{\Lambda c^2}{3}} t\right), \quad (1.62)$$

where H_{infl} is the expansion rate during the time of inflation. In the last step we used the Friedmann equation (1.19) for an Ω_Λ dominated universe. The inflationary expansion did not need to last very long, only a small period of this order of magnitude at this early in time is needed to explain both the horizon and flatness problems, as well as stretch out significant initial irregularities to create the homogeneous Universe we observe today. We assume the inflation period ended with a phase transition. During a process called reheating, the vacuum energy density was then transformed into ordinary matter and radiation, filling the Universe with radiation at roughly the same temperature as before inflation took place.

1.8.5 Observations confirming the theory

During the last decades, observations have been able to confirm and answer many important theoretical questions as well as ruling out others. Here the most important observational evidence for a flat Λ CDM universe with a hot and compact beginning are presented.

Hot Big Bang Models in which a Big Bang does not exist are constrained by a maximum redshift. Since galaxies with $z = 8.6$ have been observed (Lehnert et al. 2010) and we know that $\Omega_m > 0.1$, these models cannot be correct (see also Sect. 1.5). The abundance of light elements in the Universe further supports this theory. As the deuterium-to-hydrogen ratio is preserved throughout time, observations have been able to confirm predictions made possible only with a hot Big Bang beginning. Finally, the (hot) Big Bang theory predicts that the Universe is filled with radiation containing the remnant heat from the Big Bang. It also anticipates that this radiation has the frequency spectrum of a black-body. WMAP observations have confirmed the latter, measuring a cosmic microwave background radiation originating from the last scattering surface, having the most perfect thermal spectrum ever measured.

Accelerated expansion of the Universe As a result of studying the light curves of supernovae type Ia (SN Ia), a very tight correlation between their peak luminosity and the width of their light curves was discovered. The peak luminosity of SN Ia can therefore be measured, which, when divided by the observed flux, depends on both redshift and the cosmological model. Measuring this ratio at different redshifts thus renders a method of constraining the latter. Results from such studies yield that the Universe is expanding at an accelerated rate, hence proving the existence of dark energy, respectively a non-vanishing cosmological constant. They also show that the expansion of the Universe was decelerating at $z \gtrsim 1$. Furthermore, the abundance of galaxy clusters in the Universe as a function of redshift can be used to probe the growth of structure over cosmic time. Such structure formation studies suggest the matter and vacuum density of the Universe are $\sim 30\%$ and $\sim 70\%$ of the critical density, respectively (e.g. Mantz et al. 2010; Vikhlinin et al. 2009; Allen et al. 2008). WMAP observations of the CMB anisotropies further support this scheme (assuming reasonable values of H_0), constraining the total density of the Universe to be $\Omega_m + \Omega_\Lambda \approx 1.02 \pm 0.02$, i.e. very close to unity. They also indicate the Universe consists of 74% dark energy, 22% dark matter, and 4% ordinary matter.

Dark matter Since Zwicky’s discovery in 1933 (see Sect. 1.8.1), several other methods have been able to confirm the existence of dark matter. In addition to enabling a flat rotation curve in galaxies, dark matter has been found to have a similar effect on galaxy clusters. We also know that without dark matter, structure formation (Sect. 3.4) could not have happened as described by the concordance model, for which tiny mass fluctuations must have been present before recombination. Because of photon scattering, normal (baryonic) matter was unable to become gravitationally bound, and the fluctuations must therefore have been due to a different type of matter, not interacting with light (hence “dark”). In order to form the non-linear structures seen in our current Universe, the density contrast at the time of recombination ($z_{\text{rec}} \sim 1000$; see also Sect. 3.4.2) must

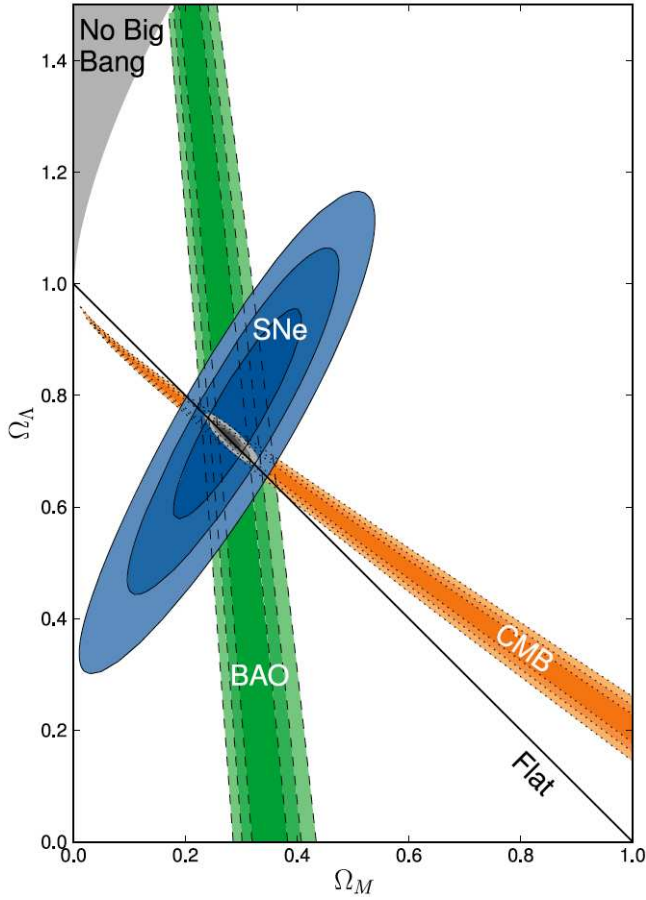


Figure 1.2: The results from SN Ia measurements, CMB anisotropy, and BAOs constraining Ω_Λ and Ω_m . The confidence regions are given on the 68.3%, 95.4%, and 99.7% levels. The solutions assume cosmological constant dark energy ($w = -1$). (Figure credit: Amanullah et al. 2010)

have been $\gtrsim 10^{-3}$. Since the CMB was released at the same time, the fluctuations in the CMB radiation should have the same amplitude – however, the CMB anisotropies render $\Delta T/T \sim 10^{-5}$. Since photons only interact with baryonic matter, this discrepancy can be accounted for by including dark matter, which may have had a higher density contrast at recombination.

Combined Using several probes of the large-scale structure of the Universe (CMB, cosmic shear, galaxy cluster/group abundance and distribution, statistical properties of the Ly α forest), density fluctuations have been measured over a large range of spatial scales. A single and smooth power spectrum fits all the various data sets very well, providing a strong confirmation for the concordance model. In addition, the results from SN Ia, CMB, and baryon acoustic oscillations (BAO; here measured in the galaxy distribution) are all in very good agreement about the Λ CDM model, as can be seen in Fig. 1.2. Acoustic oscillations arise from the encounter between gravitational attraction and gas pressure in the primordial plasma, amongst others leading to the clustering of baryonic and dark matter at scales $\sim 100 h_{70}^{-1}$ Mpc. Given the different orientations of the individual confidence regions, together the three methods put excellent constraints on Ω_Λ and Ω_m .

Chapter 2

Gravitational lensing

2.1 Introduction

Gravitational lensing is an excellent tool for studying the mass distribution in the Universe, because it does not make assumptions about the nature or the dynamical state of the gravitating matter. In particular, lensing does not discriminate between luminous and dark matter, and is thus a unique tool for studying the latter which otherwise evades observations. This aspect is very important for an analysis of the massive structures in the Universe, since dark matter dominates their state and evolution, starting from largest known objects, superclusters of galaxies, down to individual galaxies.

Gravitational lensing manifests itself in distortions of the images of background galaxies¹, whose light bundles get deflected when passing through the tidal gravitational field of an intervening mass concentration. The lensing effect can be divided into strong and weak lensing, depending on the geometry between the source, lens, and observer. Strong lensing reveals itself in multiple images of the same source, which are often highly distorted and are called arcs from the way they curve around the lens. For a massive lens, like a galaxy cluster, several independent source images can be lensed, creating an impressive system of arcs. Weak lensing, on the other hand, describes the effect lensing has on source galaxies with larger angular separation from the centre of the lens. This effect is small and only statistically measurable on a large number of source galaxies behind massive lenses. Weak lensing is therefore most commonly applied to clusters of galaxies. While strong lensing appears near the lensing centre, weak lensing can often be traced several Mpc from the centre of the galaxy cluster. In massive clusters, a combination of the two is a powerful tool to map the clusters' mass distribution.

This chapter reviews the theory behind the gravitational lensing effect, preparing the basis for the following chapters in which applications of this effect are investigated further. For thorough descriptions on gravitational lensing, the reader is referred to Munshi et al. (2008), Schneider (2006a,b), Bartelmann & Schneider (2001), Narayan & Bartelmann

¹Although easier to detect, it is not only galaxies that are lensed by foreground mass concentrations. Stars can also be lensed, e.g. as was done during the Solar eclipse of 1919 (see below for a brief summary of the lensing history). Stars lensed by stars in our own Galaxy is another example referred to as microlensing (see e.g. Wambsganss 2006). However, as this thesis mainly deals with weak lensing of galaxies, this topic will not be addressed.

(1996), and Pires et al. (2010). Note that lowercase variables written in bold are two-dimensional vectors, for which

$$\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2), \quad |\boldsymbol{\vartheta}| = \vartheta = \sqrt{\vartheta_1^2 + \vartheta_2^2} \quad (2.1)$$

hold. Here ϑ represents all lowercase variables.

2.1.1 History

Although Einstein was the first to quantitatively describe the gravitational lensing effect through General Relativity, the subject was already raised by Newton in the early 18th century. In Newtonian theory of gravitation, gravitational acceleration is independent of the mass of the falling object. Given the initial position and direction of a particle travelling in a gravitational field, its trajectory is fully determined by location and velocity. Later in the century the question arose that, if one could think of light as particles, “would they not also be affected by a gravitational field?”.

In 1915 Einstein completed his General Relativity theory (Sect 1.3). Therein the deflection angle (the angle by which light travelling in the vicinity of a spherical mass is deflected, see Sect. 2.2.1) is estimated to be twice as large as predicted by Newtonian theory. Measurements of stars deflected by the Sun during subsequent Solar eclipses proved Einstein’s theory correct.

After some time with little progression in the field, gravitational lensing was again addressed by three independent authors in the early 1960’s. Klimov (1963) studied galaxy-galaxy lensing, whilst Liebes Jr. (1964) and Refsdal (1964b,a) looked at lensing by point-mass lenses. Having significant impact in the research of gravitational lensing, their papers led to a renewed interest in the field, ultimately making way for today’s large research area of gravitational lensing.

2.2 Lensing geometry

Gravitational lens theory describes how light rays propagate in space. By making assumptions for which most astrophysically relevant situations qualify, it also allows for a much simpler explanation than that of GR (in which light propagates along the null geodesics of the space-time metric, i.e. travels along the shortest distance between two points in space-time, see Sect. 1.3). Figure 2.1 shows the geometry of a gravitational lens system in this simplified model. Light rays from a source located at a distance D_s from the observer are deflected by the lens located at a distance D_d from the observer. The distance between the source and the lens is called D_{ds} , and all distances are angular diameter distances (for this reason $D_d + D_{ds} \neq D_s$; see also Sect. 1.7). In this thesis we will only cover geometrically thin lens systems, which means that D_{ds} and D_d are much larger than the extent of the lens and the source along the line-of-sight. We can therefore treat source and lens as lying in parallel planes orthogonal to the line-of-sight. Light rays from the source are hence only affected by the lens when crossing the lens plane, otherwise they are propagating in straight lines. Since deflection angles are generally very small we can also apply the small angle approximation.

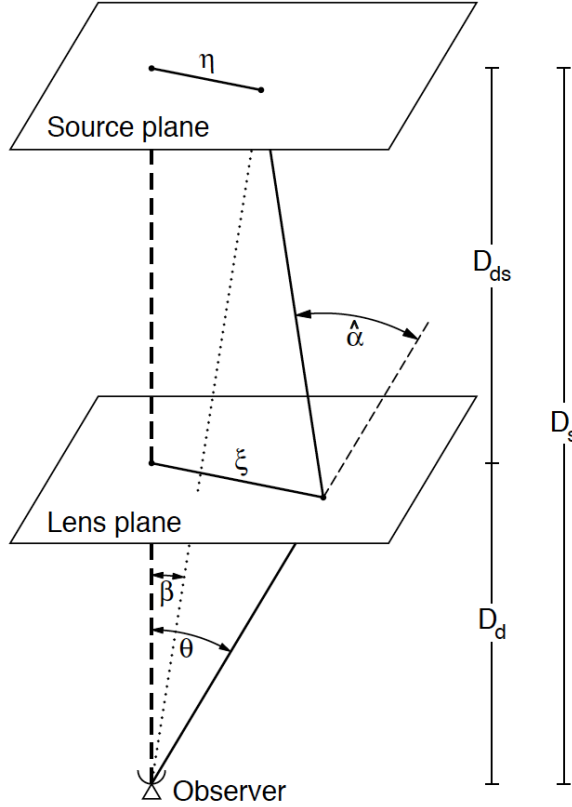


Figure 2.1: Geometry of a gravitational lens system. (Figure credit: Schneider 2006a)

2.2.1 The deflection angle

The angle, $\hat{\alpha}$, by which the light ray is deflected by the lens in Fig. 2.1 is called the deflection angle. Imagine the lens is a point-source with mass M . GR then predicts that the deflection angle of light rays with impact parameter ξ much larger than the Schwarzschild radius ($\xi \gg R_s \equiv 2GM/c^2$) is

$$\hat{\alpha} = \frac{4GM}{c^2\xi}, \quad (2.2)$$

i.e. twice the Newtonian value of the deflection angle. Here, G is the gravitational constant and c is the speed of light.

Equation (2.2) also holds for an extended radially symmetric lens with radius ξ by treating the mass inside this radius as if it was a point-mass lying at the centre of symmetry. In this special case we can treat light deflection as a one-dimensional problem by placing the origin of the coordinate system to the placement of the mock point-mass. In general, however, extended lenses are not point-masses nor are they circularly symmetric, thus the deflection angle must be treated as a two-dimensional vector in the lens plane. We know from the condition $\xi \gg R_s$ that $\hat{\alpha} \ll 1$, which also implies a small gravitational field strength $\Phi/c^2 \ll 1$, where Φ denotes the gravitational potential (galaxy clusters have typically $\Phi/c^2 \approx 0.001$ and thus satisfy this condition well). The GR field equations can

hence be linearised, and the deflection angle caused by a lens consisting of several different mass components is simply the sum of the individual deflection angles,

$$\hat{\alpha} = \sum_i \hat{\alpha}_i . \quad (2.3)$$

Consider now a three-dimensional mass distribution with volume density $\rho(\boldsymbol{\xi}, z)$, which can be divided into individual mass components of size dV and mass $dm = \rho(\boldsymbol{\xi}, z)dV$. Given this is a geometrically thin lens, we define the surface mass density of the lens at position $\boldsymbol{\xi}$ as

$$\Sigma(\boldsymbol{\xi}) \equiv \int dz \rho(\boldsymbol{\xi}, z) , \quad (2.4)$$

where $\Sigma(\boldsymbol{\xi})$ is the volume density projected onto the lens plane. Equation (2.3) can then be written

$$\hat{\alpha}(\boldsymbol{\xi}) = \int d\xi' \Sigma(\boldsymbol{\xi}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} . \quad (2.5)$$

2.2.2 The lens equation

The true position of the source in the source plane is defined by a two-dimensional vector $\boldsymbol{\eta}$. The true angular position of the source is then given by

$$\boldsymbol{\beta} = \frac{\boldsymbol{\eta}}{D_s} . \quad (2.6)$$

Correspondingly, if $\boldsymbol{\xi}$ is the position where the light ray from the source hits the lens plane, the angular position of this point would be

$$\boldsymbol{\theta} = \frac{\boldsymbol{\xi}}{D_d} , \quad (2.7)$$

i.e. the apparent angular position of the source on the sky. Taking advantage of the geometry of the lens system together with the small angle approximation, we derive

$$\begin{aligned} \boldsymbol{\eta} &= D_s \boldsymbol{\theta}(\boldsymbol{\xi}) - D_{ds} \hat{\alpha}(\boldsymbol{\xi}) \\ &= \frac{D_s}{D_d} \boldsymbol{\xi} - D_{ds} \hat{\alpha}(\boldsymbol{\xi}) . \end{aligned} \quad (2.8)$$

Upon dividing by D_s and inserting for eqs. (2.6) and (2.7) we get

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s} \hat{\alpha}(D_d \boldsymbol{\theta}) . \quad (2.9)$$

By defining the scaled deflection angle as

$$\boldsymbol{\alpha} \equiv \frac{D_{ds}}{D_s} \hat{\alpha}(D_d \boldsymbol{\theta}) , \quad (2.10)$$

we have finally derived the lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) . \quad (2.11)$$

Given the true position $\boldsymbol{\beta}$ of a source, the lens equation (2.11) yields the image position(s) $\boldsymbol{\theta}$ at which the source is observed. If, for a fixed $\boldsymbol{\beta}$, several solutions exist, the source has more than one image. The lens is then characterised as a strong lens producing multiple images.

It is useful to write the scaled deflection angle in terms of the surface mass density $\Sigma(\boldsymbol{\xi})$,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2}. \quad (2.12)$$

The dimensionless surface mass density, also called convergence, describes the isotropic focussing of light rays, and is defined as

$$\kappa(\boldsymbol{\theta}) \equiv \frac{\Sigma(D_d \boldsymbol{\theta})}{\Sigma_{\text{cr}}}, \quad (2.13)$$

where the critical surface mass density is defined as

$$\Sigma_{\text{cr}} \equiv \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{\text{ds}}}. \quad (2.14)$$

Lenses in which at least one part fulfil $\Sigma \geq \Sigma_{\text{cr}} \Leftrightarrow \kappa \geq 1$ may produce multiple images depending on the positions of their background sources. The strength of the lens can therefore be characterised in terms of κ (or Σ_{cr}), separating between weak ($\kappa \ll 1$) and strong ($\kappa \geq 1$) lenses. Although being sufficient, $\kappa \geq 1$ does not necessarily give rise to multiple images.

Using the relation $\nabla \ln |\boldsymbol{\theta}| = \boldsymbol{\theta}/|\boldsymbol{\theta}|^2$, the scaled deflection angle can be written as a gradient of the deflection potential, ψ ,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \psi(\boldsymbol{\theta}), \quad (2.15)$$

where ψ is the appropriately scaled, projected Newtonian potential of the lens, given by

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \kappa(\boldsymbol{\theta}') \ln |\boldsymbol{\theta} - \boldsymbol{\theta}'|. \quad (2.16)$$

From eq. (2.16) we can derive the Poisson equation in two dimensions,

$$\nabla_{\boldsymbol{\theta}}^2 \psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}), \quad (2.17)$$

using the fact that $\nabla^2 \ln |\boldsymbol{\theta}| = 2\pi\delta_{\text{D}}(\boldsymbol{\theta})$, where δ_{D} is the two-dimensional Dirac delta function ($\delta_{\text{D}}(\boldsymbol{x}) = +\infty$ if $\boldsymbol{x} = \mathbf{0}$, and 0 elsewhere).

2.2.3 Magnification and distortion of the lensed source image

The image shape and size of a source will change when lensed by a foreground object, due to the gravitational field through which its light bundles travel. Since gravitational light deflection neither emits nor absorbs photons, we can use Liouville's theorem to show that the surface brightness is conserved in the process. If $\boldsymbol{\theta}_0$ is a point within an image,

corresponding to the point $\beta_0 \equiv \beta(\theta_0)$ within the source, the observed surface brightness distribution in the lens plane is therefore given by

$$I(\theta) = I^s[\beta(\theta)] = I^s[\beta_0 + \mathcal{A}(\theta_0) \cdot (\theta - \theta_0)] , \quad (2.18)$$

where $I^s(\beta)$ describes the surface brightness distribution in the source plane. The last part of eq. (2.18) comes from applying the locally linearised lens equation, which is justified if the source is much smaller than the angular scale on which the lens properties change. In the lens plane $I^s(\beta)$ gets stretched along the eigenvector of the Jacobian matrix \mathcal{A} , which describes the distortion of images according to

$$\mathcal{A}(\theta) = \frac{\partial \beta}{\partial \theta} = \mathcal{I} - \frac{\partial \alpha}{\partial \theta} = \delta_{ij} - \frac{\partial^2 \psi(\theta)}{\partial \theta_i \partial \theta_j} \equiv \delta_{ij} - \psi_{,ij} . \quad (2.19)$$

Here δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$ and 0 otherwise) and \mathcal{I} is the identity matrix. The isotropic part of \mathcal{A} is given as $\delta_{ij} \text{tr} \mathcal{A} / 2$, which allows us to decompose the matrix into an isotropic and an anisotropic term,

$$\mathcal{A}(\theta) = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \quad (2.20)$$

$$= (1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - |\gamma| \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} , \quad (2.21)$$

where we have defined the convergence,

$$\kappa \equiv \frac{1}{2}(\psi_{,11} + \psi_{,22}) = \frac{1}{2} \text{tr}(\psi_{,ij}) , \quad (2.22)$$

and the complex shear, $\gamma \equiv \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}$. The components of the shear are defined as

$$\gamma_1 \equiv \frac{1}{2}(\psi_{,11} - \psi_{,22}) \equiv \gamma(\theta) \cos [2\phi(\theta)] , \quad (2.23)$$

$$\gamma_2 \equiv \psi_{,12} \equiv \psi_{,21} \equiv \gamma(\theta) \sin [2\phi(\theta)] . \quad (2.24)$$

From the above equations it is evident that convergence alone is responsible for the isotropic magnification of the source, preserving its shape but increasing its size. Shear on the other hand describes the effect of tidal gravitational forces, adding anisotropy to the lens mapping; the magnitude is given by $|\gamma| = (\gamma_1^2 + \gamma_2^2)^{1/2}$ and the orientation by ϕ . We define the observable reduced shear, $g = g_1 + ig_2$, as

$$g \equiv \frac{\gamma}{1 - \kappa} = \frac{|\gamma|}{1 - \kappa} e^{2i\phi} , \quad (2.25)$$

which enables us to rewrite the lens mapping eq. (2.20) as

$$\mathcal{A}(\theta) = (1 - \kappa) \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix} . \quad (2.26)$$

By defining

$$\kappa' \equiv 1 - \lambda(1 - \kappa) = (1 - \lambda) + \lambda\kappa , \quad (2.27)$$

$$\gamma' \equiv \lambda\gamma , \quad (2.28)$$

we can perform the invariance transformation,

$$T: \mathcal{A}(\boldsymbol{\theta}) \rightarrow \lambda \mathcal{A}(\boldsymbol{\theta}), \quad (2.29)$$

of the lens mapping,

$$\begin{aligned} \lambda \mathcal{A}(\boldsymbol{\theta}) &= \lambda(1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \lambda\gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \\ &= (1 - \kappa') \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma' \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \\ &= (1 - \kappa') \begin{pmatrix} 1 - g'_1 & -g'_2 \\ -g'_2 & 1 + g'_1 \end{pmatrix}. \end{aligned} \quad (2.30)$$

Comparing this result with eq. (2.26) and looking at the definition of the reduced shear (eq. (2.25)) we see that under this transformation g , and hence the shape of an image, does not change, i.e. $g' = g$. Using image shapes only, the surface mass density of the lens can therefore only be measured up to a constant, $(1 - \lambda)$ (see eq. (2.27)). This problem is summarised by the term ‘‘mass-sheet degeneracy’’, which can be broken by considering magnification effects or assuming that on average $\kappa = 0$ at a large distance from the lens.

Figure 2.2 illustrates the effects of convergence and shear to a circular source. The major and minor axes are given by

$$a = \frac{R_0}{1 - \kappa - |\gamma|} = \frac{R_0}{(1 - \kappa)(1 - |g|)}, \quad b = \frac{R_0}{1 - \kappa + |\gamma|} = \frac{R_0}{(1 - \kappa)(1 + |g|)}, \quad (2.31)$$

respectively, where R_0 is the intrinsic radius of the source. The magnification introduced by convergence is given by

$$\mu = \frac{1}{\det \mathcal{A}} = \frac{1}{(1 - \kappa)^2 - |\gamma|^2} = \frac{1}{(1 - \kappa)^2(1 - |g|^2)}. \quad (2.32)$$

Note that eq. (2.32) is not describing the magnification of the brightness of the source, but rather the magnification in the image size. However, due to photon number conservation this magnification also implies a corresponding increase/decrease of the source brightness.

2.2.4 Relative lensing efficiency

The strength of a gravitational lens at redshift z_d can be measured as a function of source redshift z . By comparing the convergence for a source redshift z to that of $z = \infty$,

$$\kappa(z) = \frac{\Sigma(z; z_d)}{\Sigma_{\text{cr}}(z; z_d)} = \frac{\Sigma(z; z_d)}{\Sigma_{\text{cr}}(\infty; z_d)} \frac{\Sigma_{\text{cr}}(\infty; z_d)}{\Sigma_{\text{cr}}(z; z_d)} H(z - z_d) = \kappa_\infty n(z; z_d), \quad (2.33)$$

where $\kappa_\infty = \kappa(z = \infty)$, we find that the relative lensing efficiency is defined by

$$n(z; z_d) \equiv \frac{\Sigma_{\text{cr}}(\infty; z_d)}{\Sigma_{\text{cr}}(z; z_d)} H(z - z_d). \quad (2.34)$$

Here the Heaviside step function ensures that sources in front of the lens remain unlensed ($H(x) = 0$ for $x < 0$ and 1 otherwise). Similarly we can write for the shear,

$$\gamma(z) = \gamma_\infty n(z; z_d), \quad (2.35)$$

which also preserves the relation between γ and κ given in eq. (2.96).

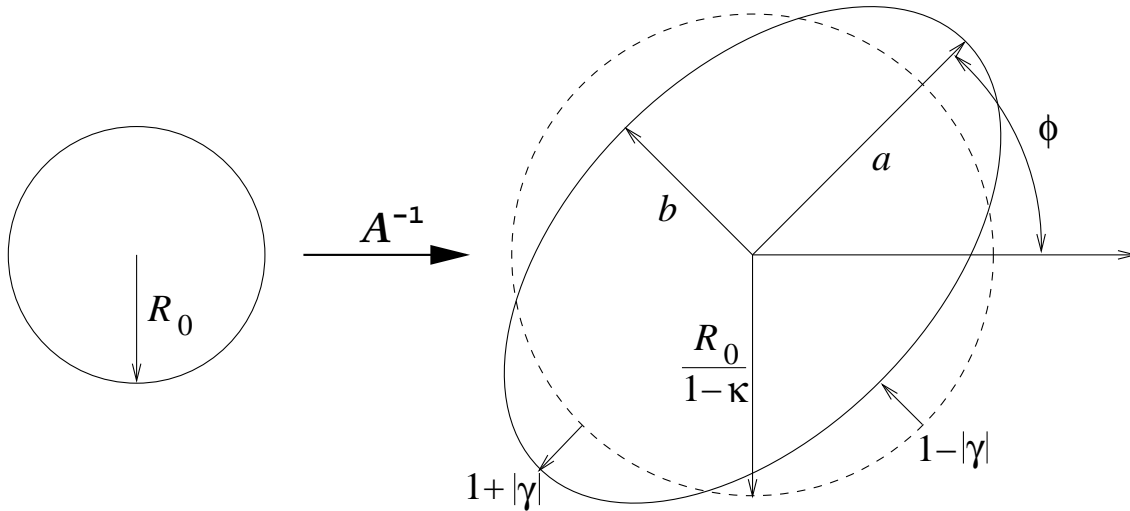


Figure 2.2: The effects of convergence and shear introduced by a lens on an intrinsically circular source. The Jacobian matrix \mathcal{A} describes the mapping between the source and image planes. (Figure credit: Schrabback 2007)

2.3 Shear measurements using weak gravitational lensing

Strong gravitational lensing includes multiple and highly distorted source images seen as arcs (see Fig. 2.3). In weak gravitational lensing the lensing effect cannot be measured individually on single galaxies but only statistically on a larger number of sources. This is inferred from the Jacobian matrix, which is very close to the unit matrix in the weak lensing regime, yielding weak distortions and small magnifications. As opposed to strong lensing, which offers detailed information about the innermost (i.e. most massive) regions of a galaxy cluster, weak lensing provides ways to measure the mass of a cluster out to very large radii, yielding the mean mass properties of the cluster and its density profile.

Weak lensing effects are detected through measurements of ellipticities of a large number of faint background galaxies. The main source of noise in weak lensing analysis is the intrinsic ellipticities of these galaxies (see eqs. (2.45) and (2.46)). In order to distinguish between distortions caused by a weak lens and the distribution of intrinsic shapes of unlensed galaxies, the ellipticities must be examined for a systematic change. In particular, a coherent tangential alignment of galaxy shapes around the cluster centre would confirm the existence of a weak lens. Section 2.3.1 details how to define and measure image ellipticities.

An additional source of error comes from the faint background galaxies being smeared by the point-spread function (PSF), caused by atmospheric turbulence and optical aberrations. The weak shear signal is hence diluted because this smearing will cause the galaxy images to appear more circular than before the smearing. In addition, PSF anisotropies distort the images, causing the galaxies to appear more elliptical, hence introducing false shear signals. Figure 2.3 shows two exposures of the same field, one taken by a ground-based observatory (right) and the other from space (left). Comparing the space-based image to the ground-based, one can clearly see the PSF effect of atmospheric smearing in the latter. In Sect. 2.3.2 we will describe how to identify and correct for all PSF effects.

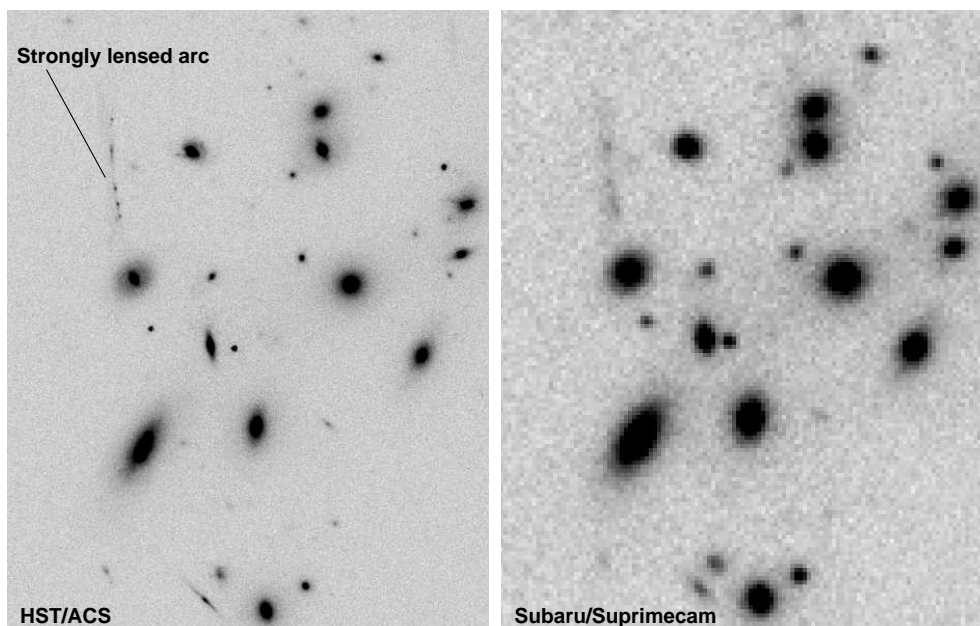


Figure 2.3: PSF effects caused by atmospheric smearing are seen in the ground-based image to the *right*, taken with Suprimecam at the Subaru telescope in Hawaii. The *left* image is taken from space using the Advanced Camera for Surveys (ACS) on board the Hubble Space Telescope (HST). Here the PSF effects are much smaller due to the absence of atmospheric dilution. Both images display the same field, and are taken with similar filters. The arcs in the images arise due to strong lensing of the cluster MS 0451-03 at redshift 0.54. The largest arc can be seen in the upper left corner and is made up of images from one multiply-imaged source.

2.3.1 Shape estimation

The measurement of weak lensing effects requires an accurate estimation of the image shapes of background galaxies, free from biases and systematics. The shear of a roughly elliptical galaxy can be estimated by the quadrupole moments of the brightness distribution as

$$Q_{ij} \equiv \frac{\int d^2\theta W_I[I(\boldsymbol{\theta})] \theta_i \theta_j I(\boldsymbol{\theta})}{\int d^2\theta W_I[I(\boldsymbol{\theta})] I(\boldsymbol{\theta})}, \quad i, j \in \{1, 2\}. \quad (2.36)$$

Here we assume $\boldsymbol{\theta}$ is measured relative to the centroid of the object's light distribution,

$$\bar{\boldsymbol{\theta}} \equiv \frac{\int d^2\theta W_I[I(\boldsymbol{\theta})] I(\boldsymbol{\theta}) \boldsymbol{\theta}}{\int d^2\theta W_I[I(\boldsymbol{\theta})] I(\boldsymbol{\theta})} = 0. \quad (2.37)$$

In both equations $W_I(I)$ is an assigned weight function. The inclusion of weights is essential, as it suppresses the noise in the data which would otherwise significantly compromise the ellipticity measurements.

The observed size and ellipticity of a galaxy image is given by the trace and traceless

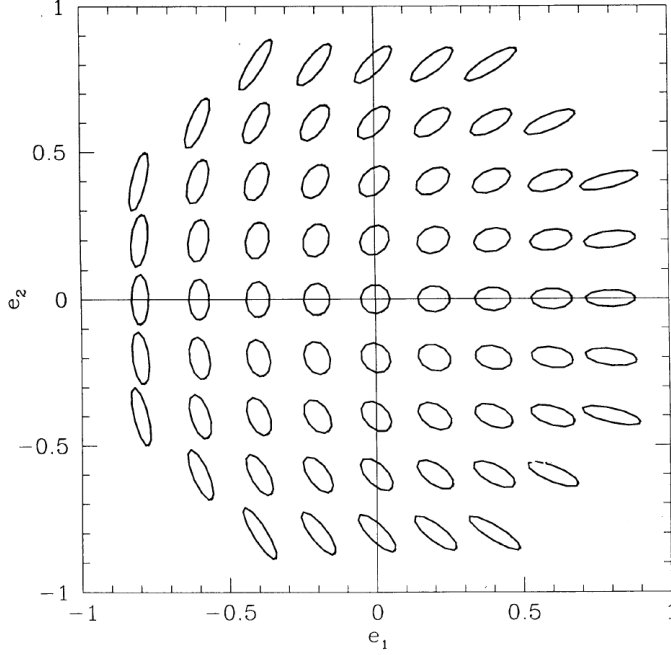


Figure 2.4: Relation between ellipticity and orientation for Gaussian ellipsoid objects. (Figure credit: Kaiser et al. 1995)

part of Q_{ij} , respectively. We can therefore define two complex ellipticities using eq. (2.36);

$$e = e_1 + ie_2 \equiv \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22}}, \quad (2.38)$$

$$\epsilon = \epsilon_1 + i\epsilon_2 \equiv \frac{Q_{11} - Q_{22} + 2iQ_{12}}{Q_{11} + Q_{22} + 2\sqrt{Q_{11}Q_{22} - Q_{12}^2}}. \quad (2.39)$$

The varying degree of ellipticity and orientation of an object will decide its placement in the $e_1 - e_2$ diagram in Fig. 2.4. The two ellipticity definitions are interchangeable; one can be obtained from the other according to

$$e = \frac{2\epsilon}{1 + |\epsilon|^2}, \quad \epsilon = \frac{e}{1 + \sqrt{1 - |e|^2}}. \quad (2.40)$$

In analogy to eqs. (2.36), (2.38), and (2.39), the second-order brightness tensor Q_{ij}^s and the complex ellipticities e^s and ϵ^s can be obtained for object shapes in the unlensed source plane. We have

$$Q_{ij}^s \equiv \frac{\int d^2\beta W_I[I^s(\beta)] \beta_i \beta_j I^s(\theta)}{\int d^2\beta W_I[I^s(\beta)] I^s(\theta)}, \quad i, j \in \{1, 2\}. \quad (2.41)$$

The transformations between source and image ellipticities have been shown to be (see e.g. Seitz & Schneider 1995; Schneider & Seitz 1995; Seitz & Schneider 1997)

$$e^s = \frac{e - 2g + g^2 e^*}{1 + |g|^2 - 2\Re[ge^*]}, \quad \epsilon^s = \begin{cases} \frac{\epsilon - g}{1 - g^* \epsilon} & |g| \leq 1, \\ \frac{1 - g\epsilon^*}{\epsilon^* - g^*} & |g| > 1, \end{cases} \quad (2.42)$$

2.3. Shear measurements using weak lensing

where the asterisk denotes the complex conjugate ($e^* = e_1 - ie_2$).

Assuming the Universe is statistically isotropic, the intrinsic ellipticities of galaxies are randomly oriented and the expectation value of the source ellipticity equals zero,

$$\langle e^s \rangle = 0 = \langle \epsilon^s \rangle . \quad (2.43)$$

In the weak lensing regime we have $\kappa \ll 1$, $|\gamma| \ll 1$, yielding $|g| \ll 1$, and the expectation value of the observed ellipticity becomes

$$\frac{\langle e \rangle}{2} \approx \langle \epsilon \rangle = g = \frac{\gamma}{1 - \kappa} \approx \gamma . \quad (2.44)$$

In other words, each image ellipticity provides an unbiased estimate of the local shear. However, this estimate is very noisy as it is subject to the intrinsic ellipticity distribution,

$$\sigma_\epsilon = \sqrt{\langle \epsilon^s \epsilon^{s*} \rangle} , \quad \epsilon \in \{e, \epsilon\} , \quad (2.45)$$

such that the 1σ rms (root mean square) error of the reduced shear is given by

$$\sigma_g = \frac{\sigma_\epsilon(1 - |g|^2)}{\sqrt{N}} \quad (2.46)$$

(Schneider et al. 2000), where N is the number of galaxy images included in the calculation, all subject to the same reduced shear. The signal-to-noise of the shear measurement is hence strongly related to the number density of background galaxies. Since the area subject to a roughly constant shear is limited, the shear must be smoothed upon averaging over galaxy images.

A coherent tangential alignment of the image ellipticities of background galaxies around a cluster centre confirms the existence of a weak lens. Comparing the tangential image ellipticities to theoretical models also enables us to infer the mass of the lens, see Sect. 2.4. The shear around the lens is measured in terms of the tangential and cross components of the shear relative to the polar angle of the separation vector between the lens and the sources as

$$\gamma_t = -\Re [\gamma e^{-2i\varphi}] , \quad \gamma_\times = -\Im [\gamma e^{-2i\varphi}] . \quad (2.47)$$

Given that the lens is located at $\boldsymbol{\theta}_0 = (\theta_{0,1}, \theta_{0,2})$ and the source images at $\boldsymbol{\theta}$, the polar angle is $\varphi = \arctan(\Delta\theta_2/\Delta\theta_1) = \arctan[(\theta_2 - \theta_{0,2})/(\theta_1 - \theta_{0,1})]$. Analogously to γ_t in eq. (2.47), the tangential component of an image ellipticity is defined as

$$\epsilon_t = -[\epsilon_1 \cos(2\varphi) + \epsilon_2 \sin(2\varphi)] = \frac{-[(\Delta\theta_1^2 - \Delta\theta_2^2)\epsilon_1 + 2\Delta\theta_1\Delta\theta_2\epsilon_2]}{\Delta\theta_1^2 + \Delta\theta_2^2} , \quad \epsilon \in \{e, \epsilon\} , \quad (2.48)$$

and the error equals that of each ellipticity component,

$$\sigma_{\epsilon_t} = \sigma_{\epsilon_1} = \sigma_{\epsilon_2} = \frac{\sigma_\epsilon}{\sqrt{2}} . \quad (2.49)$$

2.3.2 PSF corrections using the KSB+ formalism

The shapes of background sources can be measured through observations, however, they must first be corrected for PSF effects. The PSF describes an imaging system's response to a point source, and is caused by telescope effects. In the case of ground-based observations the PSF is also heavily influenced by atmospheric seeing. Seeing is defined as the full-width at half-maximum (FWHM) of the PSF, i.e. when plotting the flux of the PSF as a function of radius, the FWHM equals the width of the curve taken where the flux is half of its maximum. All PSF effects must be well understood in order to accurately correct for them and retrieve the real shapes of the galaxies. In this section a summary of the KSB+ method developed by Kaiser et al. (1995), Luppino & Kaiser (1997), and Hoekstra et al. (1998) is presented. KSB+ inverts the effects of PSF smearing and anisotropy on objects in an image, presenting a method to recover the true shear.

Since the weight $W_I(I)$ in eq. (2.36) depends on the surface brightness of the image, it is not very useful when applied to real data. KSB+ accounts for this by introducing a Gaussian distributed weight function, $W(\boldsymbol{\theta})$, such that

$$Q_{ij} \equiv \frac{\int d^2\boldsymbol{\theta} W(\boldsymbol{\theta}) \theta_i \theta_j I(\boldsymbol{\theta})}{\int d^2\boldsymbol{\theta} W(\boldsymbol{\theta}) I(\boldsymbol{\theta})}, \quad i, j \in \{1, 2\} \quad (2.50)$$

(Kaiser et al. 1995). This weight function depends on image position directly, suppressing noise at large distance from the centroid $\bar{\boldsymbol{\theta}}$. The latter is still defined according to eq. (2.37), but replacing $W_I(I)$ with $W(\boldsymbol{\theta})$. The introduction of this position-dependent weight function complicates the task of recovering the true image shapes of galaxies as affected by the interfering lens, since there is no longer a simple transformation between the brightness of the image and the source.

Ignoring the effects of measurement noise it is then possible to express the observed ellipticity of a galaxy as

$$e_\alpha = e_\alpha^s + P_{\alpha\beta}^{\text{sm}} q_\beta + P_{\alpha\beta}^\gamma g_\beta, \quad (2.51)$$

where the first term represents the intrinsic ellipticity of the galaxy, the second term the smearing of the galaxy image from the anisotropic PSF, and the third term the shift in ellipticity caused by gravitational shear (see Luppino & Kaiser 1997 with additional corrections from Hoekstra et al. 1998 for a thorough deduction of this equation). Below we will address each of these terms in turn.

The second term in eq. (2.51) describes the effects of an anisotropic PSF on the galaxy image. The smearing of object shapes caused by the seeing PSF can be explained by the convolution of a small, normalised, anisotropic PSF, $h(\boldsymbol{\theta})$, with a circularly symmetric seeing disk. As a result the quadrupole moments in eq. (2.50) will change as

$$Q'_{ij} = Q_{ij} + q_{lm} Z_{lmij}, \quad (2.52)$$

where q_{lm} defines the unweighted quadrupole moments of the PSF,

$$q_{lm} \equiv \int d^2\boldsymbol{\theta} \theta_l \theta_m h(\boldsymbol{\theta}), \quad (2.53)$$

and Z_{lmij} is given by

$$Z_{lmij} = \frac{1}{2} \int d^2\boldsymbol{\theta} I(\boldsymbol{\theta}) \frac{\partial^2 [W(\boldsymbol{\theta})\theta_i\theta_j]}{\partial\theta_l\partial\theta_m}. \quad (2.54)$$

A measure of the total PSF anisotropy is then defined by

$$q_\alpha \equiv \begin{pmatrix} q_{11} - q_{22} \\ 2q_{12} \end{pmatrix}. \quad (2.55)$$

$P_{\alpha\beta}^{\text{sm}}$ is called the smear polarisability tensor, and can be measured using

$$P_{\alpha\beta}^{\text{sm}} = X_{\alpha\beta}^{\text{sm}} - e_\alpha e_\beta^{\text{sm}}, \quad (2.56)$$

in which

$$X_{\alpha\beta}^{\text{sm}} = \frac{1}{Q_{11} + Q_{22}} \int d^2\boldsymbol{\theta} \begin{bmatrix} W + 2W'\theta^2 + W''(\theta_1^2 - \theta_2^2)^2 & 2W''(\theta_1^2 - \theta_2^2)\theta_1\theta_2 \\ 2W''(\theta_1^2 - \theta_2^2)\theta_1\theta_2 & W + 2W'\theta^2 + 4W''\theta_1^2\theta_2^2 \end{bmatrix} I(\boldsymbol{\theta}) \quad (2.57)$$

and

$$e_\alpha^{\text{sm}} = \frac{1}{Q_{11} + Q_{22}} \int d^2\boldsymbol{\theta} \begin{pmatrix} \theta_1^2 - \theta_2^2 \\ 2\theta_1\theta_2 \end{pmatrix} (2W' + W''\theta^2) I(\boldsymbol{\theta}), \quad (2.58)$$

where the prime refers to differentiation with respect to θ^2 . Because stars are unlensed foreground objects ($g_\beta = 0$) and intrinsically circular ($e_\alpha^s = 0$), applying eq. (2.51) to stellar objects provides a measure of q_α . As the changes in object shapes due to q_α are applied after the light hits Earth's atmosphere, the second term in eq. (2.51) is always the same. By modelling q_α outside the locations of the stars, i.e. at the galaxy positions, this term can be applied to all objects. Using only bright stars to determine these correction factors, the measurement noise is negligible.

The isotropic smearing by the seeing disk has a circularising effect on the ellipticities of the galaxy images, resulting in a suppression of the shear signal. This is represented by the pre-seeing shear polarisability tensor, $P_{\alpha\beta}^\gamma$, given in the third term of eq. (2.51). In order to calculate $P_{\alpha\beta}^\gamma$ we must first define the post-seeing shear polarisability tensor, $P_{\alpha\beta}^{\text{sh}}$, which can be calculated for each object together with e_α and the centroid of the object's light distribution (see eq. (2.37)). We measure the first-order shift in ellipticity caused by the shear as

$$\delta e_\alpha = P_{\alpha\beta}^{\text{sh}} \gamma_\beta. \quad (2.59)$$

Here $P_{\alpha\beta}^{\text{sh}}$ is defined as

$$P_{\alpha\beta}^{\text{sh}} = X_{\alpha\beta}^{\text{sh}} - e_\alpha e_\beta^{\text{sh}}, \quad (2.60)$$

with $X_{\alpha\beta}^{\text{sh}}$ and e_β^{sh} defined by

$$X_{\alpha\beta}^{\text{sh}} = \frac{1}{Q_{11} + Q_{22}} \int d^2\boldsymbol{\theta} \begin{bmatrix} 2W\theta^2 + 2W'(\theta_1^2 - \theta_2^2)^2 & 4W'(\theta_1^2 - \theta_2^2)\theta_1\theta_2 \\ 4W'(\theta_1^2 - \theta_2^2)\theta_1\theta_2 & 2W\theta^2 + 8W'\theta_1^2\theta_2^2 \end{bmatrix} I(\boldsymbol{\theta}) \quad (2.61)$$

and

$$e_\alpha^{\text{sh}} = 2e_\alpha + \frac{2}{Q_{11} + Q_{22}} \int d^2\boldsymbol{\theta} \begin{pmatrix} \theta_1^2 - \theta_2^2 \\ 2\theta_1\theta_2 \end{pmatrix} W'\theta^2 I(\boldsymbol{\theta}), \quad (2.62)$$

Furthermore, $P_{\alpha\beta}^\gamma$ is defined as

$$P_{\alpha\beta}^\gamma = P_{\alpha\beta}^{\text{sh}} - P_{\alpha\mu}^{\text{sm}} (P^{\text{sm}\star})_{\mu\delta}^{-1} P_{\delta\beta}^{\text{sh}\star}, \quad (2.63)$$

where the asterisk denotes $P_{\alpha\beta}^{\text{sh}}$ and $P_{\alpha\beta}^{\text{sm}}$ applied to stellar objects.

Inserting the results obtained above into eq. (2.51) yields a first-order correction for PSF anisotropies and isotropic PSF smearing, allowing for the evaluation of the reduced shear,

$$g_\beta = (P^\gamma)_{\alpha\beta}^{-1} [e_\alpha - P_{\alpha\beta}^{\text{sm}} q_\beta]. \quad (2.64)$$

Following the approach by Wold et al. (2002), we assume the PSF is close to circular after the correction, and the off-diagonal elements of the polarisability tensors in eq. (2.63) can be neglected as they are expected to be small compared to the diagonal elements. Ideally these elements would be equal to zero, though due to noise this is not the case. By also assuming the object sizes do not change when correcting the ellipticities, the description of P will remain the same after corrections. The polarisability tensors can therefore be approximated by $P_{\alpha\beta} = \frac{1}{2}(P_{11} + P_{22})\delta_{\alpha\beta}$, where $\delta_{\alpha\beta}$ denotes the Kronecker delta. The average of $P^{\text{sh}\star}/P^{\text{sm}\star}$ can then be calculated as

$$\left\langle \frac{P^{\text{sh}\star}}{P^{\text{sm}\star}} \right\rangle = \frac{1}{N_{\text{stars}}} \sum_{\text{stars}} \frac{P_{11}^{\text{sh}\star} + P_{22}^{\text{sh}\star}}{P_{11}^{\text{sm}\star} + P_{22}^{\text{sm}\star}}, \quad (2.65)$$

where using the median value rather than the mean minimises the effect of outliers. Using a scalar approximation, eq. (2.63) can therefore be written

$$p^\gamma \equiv \frac{1}{2}(P_{11}^{\text{sh}} + P_{22}^{\text{sh}}) - \frac{1}{2}(P_{11}^{\text{sm}} + P_{22}^{\text{sm}}) \left\langle \frac{P^{\text{sh}\star}}{P^{\text{sm}\star}} \right\rangle. \quad (2.66)$$

As $P_{\alpha\beta}^{\text{sh}}$ and $P_{\alpha\beta}^{\text{sm}}$ have already been calculated, this equation is easily solved, and eq. (2.64) provides an estimate of the gravitational shear of each object.

In the course of the Shear TESting Programme 2 (Massey et al. 2007a; STEP2) the performance of various shape measurement methods was evaluated based on simulations. With the KSB+ correction scheme as presented above, the true gravitational shear of lensed galaxies can be recovered with an accuracy on the order of $\sim 1\%$, which is fully sufficient to study typical galaxy clusters which have shear amplitudes of up to 10%. Better measurement algorithms are currently being developed, however, a pipeline that performs significantly and systematically better than KSB+ on real data has not yet emerged.

2.4 Lens models

2.4.1 Axially symmetric lenses

The simplest lens models assume that the mass distribution is axially symmetric. Setting its origin to the centre of symmetry yields $\Sigma(\boldsymbol{\xi}) = \Sigma(|\boldsymbol{\xi}|)$, which implies $\kappa(\boldsymbol{\theta}) = \kappa(|\boldsymbol{\theta}|)$.

This again means that the deflection angle $\alpha(\theta)$ is collinear to θ and hence also β , and the lens equation (2.11) becomes one-dimensional,

$$\beta = \theta - \alpha(\theta) . \quad (2.67)$$

The assumption of axial symmetry also renders the following properties for the deflection angle,

$$\alpha(\theta) = -\alpha(-\theta) = \frac{m(\theta)}{\theta} = \bar{\kappa}(\theta)\theta , \quad (2.68)$$

where $m(\theta)$ is the dimensionless mass and $\bar{\kappa}(\theta)$ the mean surface mass density, both inside a circle of angular radius θ ;

$$m(\theta) = 2 \int_0^\theta d\theta' \theta' \kappa(\theta') , \quad \bar{\kappa}(\theta) = \frac{m(\theta)}{\theta^2} . \quad (2.69)$$

A lens may contain critical curves, defined as smooth, closed curves for which $\det \mathcal{A} = 0$. The corresponding curves in the source plane are called caustics, found by applying the lens equation to the critical curves. Critical curves and caustics are important, one of the reasons being that the number of images produced by the lens is determined by the location of the source with respect to the caustic curves.

Applying eq. (2.32) to an axially symmetric lens yields

$$\det \mathcal{A} = (1 - \kappa)^2 - |\gamma|^2 = (1 - \bar{\kappa})(1 + \bar{\kappa} - 2\kappa) . \quad (2.70)$$

Critical curves hence occur when either $1 - \bar{\kappa}(\theta) = 0$ (tangential critical curves) or $1 + \bar{\kappa}(\theta) - 2\kappa(\theta) = 0$ (radial critical curves).

The location of the tangential critical curve enables us to calculate the mass inside it, as we can see from the following. Using eq. (2.68) we can write the lens equation (2.67) as

$$\beta = [1 - \bar{\kappa}(\theta)]\theta . \quad (2.71)$$

The tangential critical curve occurs when $\bar{\kappa}(\theta) = 1$, rendering $\beta = 0$ when inserted into the above equation. Using eqs. (2.2) and (2.10) we write

$$\beta(\theta) = \theta - \frac{D_{\text{ds}}}{D_{\text{d}}D_{\text{s}}} \frac{4GM(\theta)}{c^2\theta} , \quad (2.72)$$

which means that a source lying exactly on the optical axis is imaged as a ring given that the lens has $\Sigma > \Sigma_{\text{cr}}$. The radius of this ring, the Einstein radius, is given by setting $\beta = 0$ into the above equation,

$$\theta_{\text{E}} = \left[\frac{4GM(\theta_{\text{E}})}{c^2} \frac{D_{\text{ds}}}{D_{\text{d}}D_{\text{s}}} \right]^{1/2} . \quad (2.73)$$

Assuming that giant arcs trace the tangential critical curve, we can thus calculate the mass, $M(\theta_{\text{E}})$, inside the arc.

The Einstein radius is useful in separating between strong and weak lensing. A source with $\theta < \theta_{\text{E}}$ will experience strong lensing whereas a source at $\theta \gg \theta_{\text{E}}$ is subject to weak lensing. When a source has multiple images, the separation between the images is typically $2\theta_{\text{E}}$. Finally, we see from eqs. (2.14) and (2.73) that the mean surface mass density inside the Einstein radius equals the critical surface mass density, in other words $\bar{\kappa} = 1$.

2.4.2 The singular isothermal sphere

The mass distribution described by a singular isothermal sphere (SIS) is isothermal and spherically symmetric. Despite its simple form, this model is able to describe the lens effect in both galaxies and galaxy clusters surprisingly well.

By assuming that all mass components (e.g. stars, galaxies, etc.) behave like particles of an ideal gas, the equation of state for these components takes the form

$$p = \frac{\rho k T}{m}, \quad (2.74)$$

where ρ and m are the density and mass, respectively. The one-dimensional velocity dispersion σ_v can be related to the temperature T through

$$m\sigma_v^2 = kT \quad (2.75)$$

due to the assumption of thermal equilibrium. In an isothermal gas the temperature, thus also the velocity dispersion, is constant across the distribution of mass. We therefore utilise the equation of hydrostatic equilibrium to write

$$\rho(r) = \frac{\sigma_v^2}{2\pi G} \frac{1}{r^2}. \quad (2.76)$$

Since $\rho \propto r^{-2}$, the density diverges for $r \rightarrow 0$, introducing the need to define a finite core radius. In addition, as we will see shortly, the mass diverges as $r \rightarrow \infty$ such that $M(r) \propto r$. A radial cut-off must therefore be made in order to find the total mass of the distribution.

By projecting eq. (2.76) along the line-of-sight we find the surface mass density of an SIS lens;

$$\Sigma(\xi) = \int_{-\infty}^{\infty} dz \rho(\sqrt{\xi^2 + z^2}) = \frac{\sigma_v^2}{2G\xi}, \quad (2.77)$$

from which we can calculate the projected mass $M(\xi)$ inside the radius ξ

$$M(\xi) = 2\pi \int_0^\xi d\xi' \xi' \Sigma(\xi') = \frac{\pi\sigma_v^2\xi}{G}. \quad (2.78)$$

Using eqs. (2.2) and (2.10) and substituting for eq. (2.78) the deflection angle of an SIS is found. However, looking at eq. (2.73) we see that the Einstein angle equals that of α , i.e.

$$\alpha(\theta) = \frac{D_{\text{ds}}}{D_{\text{s}}} \hat{\alpha}(D_{\text{d}}\theta) = 4\pi \left(\frac{\sigma_v}{c}\right)^2 \frac{D_{\text{ds}}}{D_{\text{s}}} \equiv \theta_{\text{E}}. \quad (2.79)$$

We can therefore write

$$\kappa(\theta) = |\gamma|(\theta) = \frac{\theta_{\text{E}}}{2|\theta|}, \quad \bar{\kappa}(\theta) = \frac{\theta_{\text{E}}}{|\theta|}, \quad \alpha(\theta) = \theta_{\text{E}} \frac{\theta}{|\theta|}. \quad (2.80)$$

The lens can only produce multiple images of a source that lies within the Einstein radius, $\beta < \theta_{\text{E}}$. The lens equation then has two solutions

$$\theta_{\pm} = \beta \pm \theta_{\text{E}}. \quad (2.81)$$

Due to the symmetry of the SIS model, the lens, source, and images all lie on a straight line.

The magnification of a circularly symmetric lens is given by

$$\mu = \frac{\theta}{\beta} \frac{d\theta}{d\beta}. \quad (2.82)$$

By applying eq. (2.81) we then find the magnifications of the two images

$$\mu = \frac{\theta_{\pm}}{\beta} = 1 \pm \frac{\theta_E}{\beta} = \left(1 \mp \frac{\theta_E}{\theta_{\pm}}\right)^{-1}. \quad (2.83)$$

If the source lies outside the Einstein radius, $\beta > \theta_E$, only one image will be seen at $\theta = \theta_+ = \beta + \theta_E$.

2.4.3 The Navarro, Frenk, & White profile

The Navarro, Frenk, & White (1997, 1995; NFW) profile is derived from fitting the density profiles of numerically simulated cold dark matter halos. It appears to give a very good description of the radial mass distribution inside the virial radius (see Sect. 3.3.1) of a galaxy cluster (Wright & Brainerd 2000).

The general NFW density profile follows

$$\rho(r) = \frac{\delta_{\text{cr}} \rho_{\text{cr}}(z)}{(r/r_s)(1+r/r_s)^2}, \quad (2.84)$$

where δ_{cr} is the characteristic over-density of the halo defined as

$$\delta_{\text{cr}} \equiv \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)}, \quad (2.85)$$

and $\rho_{\text{cr}}(z)$ is the critical density of the Universe at redshift z of the halo (i.e. the generalised version of eq. (1.31)). The characteristic radius of the halo, the scale radius r_s , is defined in terms of the concentration parameter,

$$c \equiv \frac{r_{200}}{r_s}, \quad (2.86)$$

where r_{200} denotes the radius inside which the mean mass density of the halo equals $200\rho_{\text{cr}}(z)$. The concentration of mass towards the inner regions of a halo increases with c . For $r \ll r_s$ the density profile follows $\rho \propto r^{-1}$, whereas for $r \gg r_s$ it follows $\rho \propto r^{-3}$. The scale radius therefore defines the radius at which the slope of the density profile changes. The mass of an NFW halo inside r_{200} is given by

$$M_{200} \equiv M(r_{200}) = \frac{800\pi}{3} \rho_{\text{cr}} r_{200}^3. \quad (2.87)$$

The deflection potential of an NFW lens is defined in terms of its three-dimensional potential, $\Phi(D_d \boldsymbol{\theta}, z)$, as

$$\psi(\boldsymbol{\theta}) \equiv \frac{D_{\text{ds}}}{D_d D_s} \frac{2}{c^2} \int dz \Phi(D_d \boldsymbol{\theta}, z) \quad (2.88)$$

(see also Sect. 2.2.2), from which the convergence and shear can be derived using eqs. (2.22), (2.23), and (2.24). Here θ is the angular radius of the lens.

The radial dependence of the surface mass density of a spherically symmetric lens is given by integrating the three-dimensional density profile in eq. (2.4) along the line-of-sight. Inserting for eq. (2.84) and introducing $x = \xi/r_s$ results in a corresponding expression for an NFW lens (e.g. Bartelmann 1996),

$$\Sigma_{\text{NFW}}(x) = \begin{cases} \frac{2r_s\delta_{\text{cr}}\rho_{\text{cr}}}{x^2 - 1} \left[1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh}\sqrt{\frac{1-x}{1+x}} \right] & x < 1, \\ \frac{2r_s\delta_{\text{cr}}\rho_{\text{cr}}}{3} & x = 1, \\ \frac{2r_s\delta_{\text{cr}}\rho_{\text{cr}}}{x^2 - 1} \left[1 - \frac{2}{\sqrt{x^2-1}} \operatorname{arctan}\sqrt{\frac{x-1}{1+x}} \right] & x > 1. \end{cases} \quad (2.89)$$

Using eq. (2.13), the radial dependence of the convergence is given as

$$\kappa_{\text{NFW}}(x) = \frac{\Sigma_{\text{NFW}}(x)}{\Sigma_{\text{cr}}}. \quad (2.90)$$

The radial dependence of the shear can be derived using the fact that the NFW density profile is spherically symmetric,

$$\gamma_{\text{NFW}}(x) = \frac{\bar{\Sigma}_{\text{NFW}}(x) - \Sigma_{\text{NFW}}(x)}{\Sigma_{\text{cr}}} \quad (2.91)$$

(Miralda-Escude 1991), where $\bar{\Sigma}_{\text{NFW}}(x)$ denotes the mean surface mass density of an NFW halo,

$$\begin{aligned} \bar{\Sigma}_{\text{NFW}}(x) &= \frac{2}{x^2} \int_0^x dx' x' \Sigma_{\text{NFW}}(x') \\ &= \begin{cases} \frac{4r_s\delta_{\text{cr}}\rho_{\text{cr}}}{x^2} \left[\frac{2}{\sqrt{1-x^2}} \operatorname{arctanh}\sqrt{\frac{1-x}{1+x}} + \ln\left(\frac{x}{2}\right) \right] & x < 1, \\ 4r_s\delta_{\text{cr}}\rho_{\text{cr}} \left[1 + \ln\left(\frac{1}{2}\right) \right] & x = 1, \\ \frac{4r_s\delta_{\text{cr}}\rho_{\text{cr}}}{x^2} \left[\frac{2}{\sqrt{x^2-1}} \operatorname{arctan}\sqrt{\frac{x-1}{1+x}} + \ln\left(\frac{x}{2}\right) \right] & x > 1. \end{cases} \end{aligned} \quad (2.92)$$

The radial dependence of the shear is then written as (Bartelmann 1996)

$$\gamma_{\text{NFW}}(x) = \begin{cases} \frac{r_s\delta_{\text{cr}}\rho_{\text{cr}}}{\Sigma_{\text{cr}}} g_{<}(x) & x < 1, \\ \frac{r_s\delta_{\text{cr}}\rho_{\text{cr}}}{\Sigma_{\text{cr}}} \left[\frac{10}{3} + 4 \ln\left(\frac{1}{2}\right) \right] & x = 1, \\ \frac{r_s\delta_{\text{cr}}\rho_{\text{cr}}}{\Sigma_{\text{cr}}} g_{>}(x) & x > 1, \end{cases} \quad (2.93)$$

where the functions $g_{</>}(x)$ are independent of the cosmology, depending only on the dimensionless radius x ,

$$g_{<}(x) = \frac{8 \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}}}{x^2 \sqrt{1-x^2}} + \frac{4}{x^2} \ln \left(\frac{x}{2} \right) - \frac{2}{x^2-1} + \frac{4 \operatorname{arctanh} \sqrt{\frac{1-x}{1+x}}}{(x^2-1) \sqrt{1-x^2}}, \quad (2.94)$$

$$g_{>}(x) = \frac{8 \operatorname{arctan} \sqrt{\frac{x-1}{1+x}}}{x^2 \sqrt{x^2-1}} + \frac{4}{x^2} \ln \left(\frac{x}{2} \right) - \frac{2}{x^2-1} + \frac{4 \operatorname{arctan} \sqrt{\frac{x-1}{1+x}}}{(x^2-1)^{3/2}}. \quad (2.95)$$

2.5 Two-dimensional mass reconstruction

Kaiser & Squires (1993) discovered that distorted images of background galaxies can be used for parameter-free reconstructions of surface mass densities of galaxy clusters. Their method will be explained in this section.

Inserting eq. (2.16) into eqs. (2.23) and (2.24) yields a relation between the complex shear and the convergence as

$$\gamma(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \kappa(\boldsymbol{\theta}') \mathcal{D}(\boldsymbol{\theta} - \boldsymbol{\theta}'), \quad (2.96)$$

showing that the shear is a convolution of the surface mass density with the kernel

$$\mathcal{D}(\boldsymbol{\theta}) \equiv \frac{-\theta_1^2 + \theta_2^2 - 2i\theta_1\theta_2}{\theta^4}. \quad (2.97)$$

The Fourier transform of the surface mass density and its inverse are defined as

$$\hat{\kappa}(\boldsymbol{l}) \equiv \int_{\mathbb{R}^2} d^2\theta \kappa(\boldsymbol{\theta}) \exp(i\boldsymbol{l} \cdot \boldsymbol{\theta}), \quad (2.98)$$

$$\kappa(\boldsymbol{\theta}) \equiv \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} d^2l \hat{\kappa}(\boldsymbol{l}) \exp(-i\boldsymbol{l} \cdot \boldsymbol{\theta}), \quad (2.99)$$

where \boldsymbol{l} is the two-dimensional angular Fourier variable. Fourier transforming eqs. (2.22), (2.23), and (2.24) and inserting for $\hat{\psi} = 2\hat{\kappa}/(-l^2)$ in the last two equations yields

$$\hat{\gamma} = \frac{(l_1^2 - l_2^2 + 2il_1l_2)}{l^2} \hat{\kappa}, \quad (2.100)$$

and hence also

$$\hat{\kappa} = \frac{(l_1^2 - l_2^2)\hat{\gamma}_1 + 2il_1l_2\hat{\gamma}_2}{l^2}. \quad (2.101)$$

Using the convolution theorem, the Fourier transform of eq. (2.96) becomes a multiplication rather than a convolution between κ and the kernel \mathcal{D}

$$\hat{\gamma}(\boldsymbol{l}) = \frac{1}{\pi} \hat{\kappa}(\boldsymbol{l}) \hat{\mathcal{D}}(\boldsymbol{l}) \quad \text{for } \boldsymbol{l} \neq \mathbf{0}. \quad (2.102)$$

Together with eqs. (2.100) and (2.101), this equation renders an expression for the Fourier transform of \mathcal{D} ,

$$\hat{\mathcal{D}}(\boldsymbol{l}) = \pi \frac{(l_1^2 - l_2^2 + 2il_1l_2)}{l^2}. \quad (2.103)$$

Since we know that $\hat{\mathcal{D}}(\boldsymbol{l})\hat{\mathcal{D}}^*(\boldsymbol{l}) = \pi^2$ (where the asterisk denotes complex conjugation), eq. (2.103) can be inverted to render

$$\hat{\kappa}(\boldsymbol{l}) = \frac{1}{\pi}\hat{\gamma}(\boldsymbol{l})\hat{\mathcal{D}}^*(\boldsymbol{l}) \quad \text{for } \boldsymbol{l} \neq \mathbf{0} . \quad (2.104)$$

Using the convolution theorem once more to Fourier back-transform this equation leads to the final result,

$$\begin{aligned} \kappa(\boldsymbol{\theta}) - \kappa_0 &= \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \mathcal{D}^*(\boldsymbol{\theta} - \boldsymbol{\theta}') \gamma(\boldsymbol{\theta}') \\ &= \frac{1}{\pi} \int_{\mathbb{R}^2} d^2\theta' \Re [\mathcal{D}^*(\boldsymbol{\theta} - \boldsymbol{\theta}') \gamma(\boldsymbol{\theta}')] , \end{aligned} \quad (2.105)$$

thus providing a simple method for inferring the mass distribution of a galaxy cluster. In the last step of eq. (2.105) we have utilised the fact that the surface mass density must be real. The constant κ_0 arises due to $\hat{\mathcal{D}}$ being undefined for $\boldsymbol{l} = \mathbf{0}$, which means that the convergence can be determined from the shear only up to an additive constant. This makes sense also physically, since a uniform surface mass density does not provide any shear. This mass-sheet degeneracy is a more constrained case as compared to the mass-sheet degeneracy described in Sect. 2.2.3, as the convergence in eq. (2.105) is known up to a single additive constant. Note that this method suffers from significant boundary effects since it assumes an infinitely large data field.

Chapter 3

Clusters of galaxies

A profound understanding of the matter distribution in the Universe is one of the leading questions in today's cosmology. Being the most massive bound and relaxed structures in the Universe, clusters of galaxies signify the most prominent density peaks of the large-scale structure. Since the time it takes the cluster mass distribution to settle into equilibrium (Sect. 3.3.1) is comparable to or larger than the age of the Universe, the evolution of clusters provides a direct insight into the growth of cosmic structures. Structure formation comprises matter density fluctuations forming increasingly larger structures due to gravitational forces, in which galaxy clusters are the largest objects that have had time to collapse under the influence of their own gravitation.

Having a high galaxy density, clusters also provide ideal environments for analysing interactions between galaxies, as well as the local galaxy density's impact on the morphology and evolution of galaxies. Because the gravitational potential wells of clusters are so deep, clusters retain all their gaseous matter, despite the enormous energy input connected to supernovae and active galactic nuclei. A wealth of information about galaxy formation can therefore be extracted from the baryonic component of clusters.

Galaxy clusters comprise one of the main cosmological probes. As shown by the Virgo Consortium on representatively large N-body simulations (see e.g. Evrard et al. 2002), the statistical distribution of clusters is a strong function of mass and redshift, and their evolution depends critically on the background cosmology.

In this chapter we present an overview of galaxy clusters, which are the main research topic of this thesis (see Chaps. 5, 6, and 7). More extensive information about clusters can be found in Schneider (2006c), Voit (2005), and Borgani & Guzzo (2001). A summary of the Navarro, Frenk, & White (1997) density profile can be found in Wright & Brainerd (2000). Additionally, Comins & Kaufmann (2003) and the free encyclopedia Wikipedia¹ provide general information about most topics in this chapter.

3.1 Optical clusters

Containing most of the visible matter in the Universe, galaxies consist of an assembly of gravitationally bound stars, gas, and dust. Although the visible Universe consists

¹<http://en.wikipedia.org/>

of $\sim 1.3 \times 10^{11}$ galaxies (extrapolated from the Hubble Ultra Deep Field), their overall shapes show surprising consistency. In the 1920s Hubble began cataloguing galaxies, classifying them according to his tuning fork diagram shown in Fig. 3.1 (Hubble 1936). Elliptical galaxies are classified according to their ovalness, going from circular (E0) to significantly elongated (E7). They contain relatively little molecular gas and dust in which star formation takes place. Elliptical galaxies thus have a very low star-formation rate, and are mainly comprised of low-mass, long-lived stars, giving them a redder colour than spiral galaxies. S0 or SB0 galaxies, also called lenticulars, are intermediate types between ellipticals and spirals. They are similar to spiral galaxies, having both a central bulge and a disk, though lacking the spiral arms. Ellipticals and lenticulars are sometimes called early-type galaxies due to their placement in Hubble’s tuning fork diagram. Due to their old stellar populations, early-type galaxies form a relatively well-defined colour-magnitude sequence at any given redshift. Such galaxies are commonly found in galaxy clusters, where they form a distinct “cluster red sequence” in a colour-magnitude diagram comprising all the members galaxies (see e.g. Fig. 5.4).

Spiral and barred spiral galaxies are classified according to the size of their central bulge and how tightly their arms wind around the centre (the two are correlated). We see that the tightness of the spiral and the size of the nuclear bulge is decreasing when going from left to right in Fig. 3.1. The amount of gas and dust, and hence also star formation, is increasing correspondingly, giving spirals a blue colour.

In addition to the categories described above, a separate galaxy type exists, not fitting any of the classifications given. Irregular galaxies have no definite structure, are generally rich in interstellar gas and dust, and comprise both young and old stars. They are typically smaller and less massive than spiral galaxies, and increasingly more common when going to higher redshift (i.e. with decreasing age of the observable Universe). The irregular galaxies that have almost no organised structure and several OB associations (groups of very young, massive stars) are denoted Irr I. The other group of irregulars is called Irr II and encompasses highly distorted, completely asymmetrical galaxies, thought to be created by collisions between galaxies or by violent activity in their nuclei. Spiral and irregular galaxies are denoted late-type galaxies. The Hubble classification of galaxies is still in use today.

Rather than being randomly distributed throughout the Universe, galaxies assemble in groups and clusters of galaxies. Belonging to the largest known gravitationally bound structures, they form the densest part of the large-scale structure. While groups contain from only a few up to a few tens of galaxies within a sphere of diameter $d \lesssim 2.0 h_{70}^{-1}$ Mpc, clusters contain up to as much as a few thousand within diameters of $d \gtrsim 2.0 h_{70}^{-1}$ Mpc. The mass range of groups and clusters is $10^{12} M_{\odot} \lesssim M \lesssim 10^{15} M_{\odot}$. There is no sharp dividing line between groups and clusters, and we will therefore refer to galaxy clusters as comprising both terms for the remainder of this chapter.

Galaxy clusters themselves group in superclusters, who are gravitationally unbound if their mean density drops below $\sim 2.4 \rho / \rho_{\text{cr}}$ (Araya-Melo et al. 2009). Here ρ_{cr} is the critical density of the Universe, given in eq. (1.31). Large voids can be found between superclusters, in which very few galaxies exist. Being roughly spherical these voids are 30 – 120 Mpc in diameter. Most galaxies are found to be distributed on the surfaces between voids.

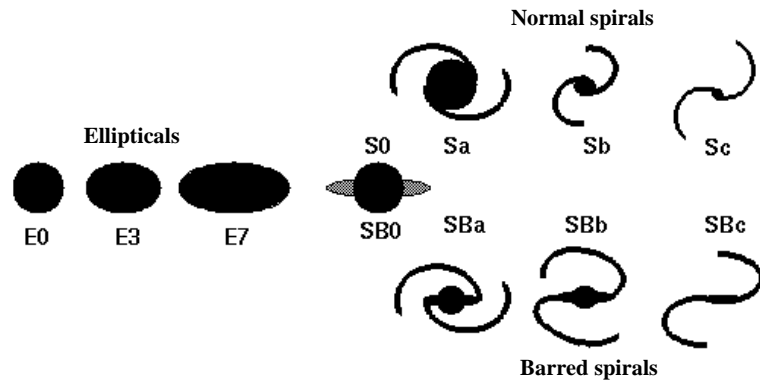


Figure 3.1: Hubble’s tuning fork diagram, summarising his galaxy classification scheme. See the text for details. (Figure credit: adapted from <http://www.astro.virginia.edu/~jh8h/glossary/hubbleclass.htm>)

3.1.1 History

While cataloguing galaxies (at the time known as nebulae), both Messier (1784) and Herschel (1785) independently recognised concentrations of galaxies on the sky while trying to determine the relation between nebulae and our own galaxy, the Milky Way. These observations indicate the first discoveries of galaxy clusters. In the beginning of the 20th century it became clear that nebulae are galaxies themselves (e.g. Hubble 1929a,b; see also Sect. 1.1), leading to an increased interest in galaxy clusters as physical systems. In 1933 Zwicky discovered the need for dark matter in clusters (see Sect. 1.8.1). Other studies looked at the evolution of galaxies within the cluster framework (see e.g. Spitzer & Baade 1951). In the 1950’s studies of galaxy clusters started covering all aspects, and had by the end of the century become one of the main research topics in extragalactic astrophysics.

Abell published his first catalogue of galaxy clusters in 1958, a catalogue that has remained one of the most important since it was published, and is still in use today. Its homogeneous selection of 2712 rich galaxy clusters is remarkable, especially considering that Abell manually searched for clusters using photographic plates. Looking at the projected overdensities of galaxies he subjectively chose a radius we now know is close to the cluster’s virial radius (see Sect. 3.3.1). His other selection criteria were also very well chosen, resulting in a final catalogue of clusters containing ≥ 50 galaxies within a magnitude interval $m_3 \leq m \leq m_3 + 2$, where m_3 denotes the apparent magnitude (Sect. 4.3.1) of the third brightest cluster galaxy. Providing the community with a statistical homogeneous cluster list, researchers could begin to study clusters as a population rather than individual objects.

3.1.2 Morphological classification

Clusters of galaxies are characterised as either regular or irregular clusters. Regular clusters appear spherically symmetric, comprising mainly early-type galaxies, with a high concentration of cluster members towards the cluster centre. They are considered to be dynamically relaxed. The richest and most massive clusters in the Universe are regular clusters. On the opposite side we find irregular clusters, having little or no symmetry, and

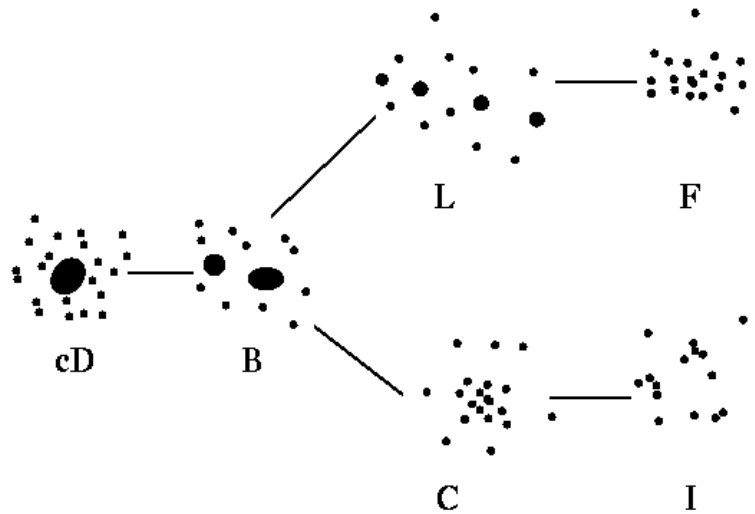


Figure 3.2: The Rood-Sastry classification; a rough morphological classification of clusters of galaxies. The concentration of the clusters decreases from left to right in the tuning-fork, whereas evolution runs from right to left. See the text for more details. (Figure credit: <http://www.astr.ua.edu/keel/galaxies/clusters.html>)

with no significant central concentration of cluster members but rather several smaller subclusters. Their galaxy population comprise all types of galaxies, in which the fraction of spiral galaxies can become as high as that of field galaxies. Irregular clusters are still in an evolutionary state.

Several, more detailed, classification systems exist. Using the Abell clusters, Rood & Sastry (1971) developed a classification scheme based on the morphology and projected galaxy distribution of the ten brightest members of the cluster. The Rood-Sastry classification divides clusters into 6 subcategories:

cD The cluster is dominated by a central cD galaxy. A cD galaxy is also known as a supergiant elliptical – it is very large and has an extended low surface brightness stellar envelope.

B The cluster is dominated by a pair of luminous galaxies in the centre.

L At least three of the ten brightest galaxies in the cluster are close to linearly aligned.

C Four or more galaxies form a single cluster core.

F The cluster’s brightest galaxies form an oblate galaxy distribution.

I The galaxy distribution is irregular, with no obvious centre or core.

As seen in Fig. 3.2 the different cluster categories appear to take on a tuning-fork pattern, in which the concentration decreases from left to right. The cluster types cD, B, L, and C are characterised as regular clusters, whereas F and I are irregular clusters. The three types in the middle, L, C, and F, are considered intermediate, in between regular and irregular. An evolution can be considered running from right to left in the diagram.

3.2 X-ray clusters

Galaxy clusters are amongst the brightest extragalactic X-ray sources in the Universe. Their X-ray emission does not originate in individual galaxies but is spatially extended. Clusters have deep potential wells which compress the baryonic gas present and heat it to X-ray emitting temperatures. The depth of a cluster's potential well can therefore be inferred from its gas temperature, which is determined through fits to the cluster's X-ray spectrum. The mass of a galaxy cluster can also be obtained from its X-ray temperature, see Sect. 3.3 for more details. Massive clusters have gas temperatures of typically $k_{\text{B}}T \sim 5 \text{ keV}$, where k_{B} is the Boltzmann constant. The spectral energy distribution of the X-rays tells us that the emission process is optically thin thermal bremsstrahlung from hot gas. In other words, X-ray radiation is created by electrons being accelerated in the electric field around protons and atomic nuclei. Since the metallicity of this gas is $\sim 1/3$ of the metallicity of the Sun, it can be inferred that a significant fraction of this intergalactic gas (the intracluster medium; ICM), was initially incorporated in galaxies. It is ejected into the ICM either through ram-pressure stripping while galaxies fall into the cluster, through the explosions of supernovae, or from jets of active galactic nuclei.

The magnitudes of the emission lines in the X-ray spectrum may indicate the abundances of elements in the intracluster medium, like iron, oxygen, and silicon. The hotter the gas in a cluster is, the higher its ionisation, and the weaker its line emission is. The emissivity for bremsstrahlung is $\propto T^{1/2}$, however the total X-ray emission of clusters increases towards lower temperatures as line emission becomes more important. Since the emissivity depends quadratically on the density of the plasma, both the spatial distribution and the X-ray luminosity of the gas is needed in order to estimate the mass of a galaxy cluster from its X-ray emission. Nevertheless, since clusters satisfy several scaling relations (i.e. strong correlations between important physical properties), an empirical relation between mass and X-ray luminosity can be utilised (see Sect. 3.3).

Clusters of galaxies were originally selected using optical methods, looking for spatial overdensities of galaxies on the sky, see Sect. 3.1.1. This method is limited by the superposition of foreground and background objects, which may lead to coincidental overdensities in the projected galaxy distribution, mocking real galaxy clusters. With X-ray telescopes, a more reliable way of selecting clusters could be developed. The high gas temperature indicates a deep potential well, hence an overdensity of matter, implying negligible projection effects. The selection of clusters is actually based on X-ray luminosity, not temperature. However, since the relation between temperature and luminosity is fairly close, a cluster's luminosity is also a good proxy for its mass (see end of Sect. 3.3).

Since the low-energy X-ray emission lines disappear with increasing cluster temperature, it is difficult to determine the redshift of a cluster from X-ray spectra alone. The emission lines of a hot cluster (e.g. the iron K line at $\sim 6 \text{ keV}$) are not very strong – poor photon statistics is thus the primary limiting factor in such redshift determination methods. Furthermore, the CCD resolution is fairly poor in this energy range ($\gtrsim 3 \text{ keV}$), resulting in a broadening of the X-ray emission lines, making it difficult to pinpoint them at specific wavelengths. X-ray clusters are therefore usually followed up by optical observations from which their redshifts are estimated – either from averaging over photometric redshifts of many cluster members, or from spectroscopic observations of the brightest cluster galaxy in addition to a few selected cluster members (Sect. 4.4).

3.2.1 Morphological classification

Using the surface brightness of its X-ray emission, a cluster can be roughly classified as either regular or irregular – the same classification scheme as was used in Sect. 3.1.2 but then based on the cluster’s galaxy distribution. Regular clusters have a smooth brightness distribution which is centred on its optical centre and decreases with radius. They typically have high X-ray luminosities and temperatures. Irregular clusters have several brightness maxima who are usually centred on single or subgroups of cluster galaxies. They may also have high temperatures due to their merger processes, where the gas is heated by shock fronts. The X-ray luminosity and temperature of a cluster normally decreases with increasing fraction of spiral galaxies.

3.3 Measuring the cluster mass

Weak gravitational lensing is the main method used throughout this thesis for determining the masses of galaxy clusters (see Chaps. 5, 6, and 7). In the following we introduce two other methods that are commonly used to infer the mass of a galaxy cluster.

3.3.1 Virial analysis

A virial analysis assumes that the cluster is in virial equilibrium, comprising a galaxy distribution similar to that of the total mass, and with an isotropic velocity dispersion. A cluster reaches virial equilibrium when the velocities of its member galaxies become randomised, commencing a pressure balance with the cluster’s gravitational forces, and assuming a configuration whose properties do not undergo any further change. The virial theorem thus applies,

$$2E_{\text{kin}} + E_{\text{pot}} = 0 , \quad (3.1)$$

where

$$E_{\text{kin}} = \frac{1}{2} \sum_i m_i v_i^2 \quad \text{and} \quad E_{\text{pot}} = -\frac{1}{2} \sum_{i \neq j} \frac{G m_i m_j}{r_{ij}} \quad (3.2)$$

represent the kinetic and the potential energy of the cluster members, respectively. The mass of the i -th galaxy is given by m_i , the absolute value of its velocity by v_i , the spatial separation between the i -th and the j -th galaxy is given by r_{ij} , and G denotes the gravitational constant. The total kinetic energy of the cluster can be written as

$$E_{\text{kin}} = \frac{1}{2} M \sigma^2 , \quad (3.3)$$

where the total mass of the cluster is defined as $M \equiv \sum_i m_i$ and the velocity dispersion as $\sigma^2 \equiv \frac{1}{M} \sum_i m_i v_i^2$. Assuming the cluster is spherical with virial radius

$$r_{\text{vir}} \equiv 2M^2 \left(\sum_{i \neq j} \frac{m_i m_j}{r_{ij}} \right)^{-1} , \quad (3.4)$$

its total potential energy is given by

$$E_{\text{pot}} = -\frac{GM^2}{r_{\text{vir}}} . \quad (3.5)$$

The virial radius is defined as the three-dimensional radius within which the matter of the cluster is virialised. Galaxies outside r_{vir} have yet to fall towards the centre of the cluster for the first time, after which they dynamically cool down into the cluster centre through dynamical friction (Chandrasekhar 1943). Dynamical friction is the process in which a particle (here galaxy) is moving through a field of massive particles (other galaxies). Due to the attraction from the moving particle's gravitational field, the massive particles will start moving towards its trajectory, creating an overdensity of matter behind it, hence gradually slowing it down.

Until now we have only considered three-dimensional quantities. However, as we can only observe projections on the sky, the above equations must be transformed into two dimensions. First we will look at how the redshift of a cluster galaxy can be converted into its line-of-sight velocity. By measuring the redshifts of all the cluster galaxies, $z_{i,\text{tot}}$, we can find the mean redshift of the cluster, $z_{\text{cl}} = \langle z_{i,\text{tot}} \rangle$, corresponding to the cluster's mean motion due to the expansion of the Universe (see e.g. Sect. 1.1). Corresponding to the individual velocities in the cluster potential, each galaxy has some additional, very small redshift component, z_i , which is calculated using the relation

$$1 + z_{i,\text{tot}} = (1 + z_{\text{cl}}) \cdot (1 + z_i) \quad (3.6)$$

(Harrison 1974; Danese et al. 1980). We then substitute for the line-of-sight velocity, $V = cz$, and solve for the individual galaxy velocity,

$$V_i = \frac{V_{i,\text{tot}} - V_{\text{cl}}}{1 + (V_{\text{cl}}/c)} = c \left(\frac{z_{i,\text{tot}} - z_{\text{cl}}}{1 + z_{\text{cl}}} \right) . \quad (3.7)$$

In other words, the observed redshift dispersion must be corrected for the expansion of the Universe. The line-of-sight velocity dispersion of the cluster can then be calculated according to

$$\sigma_V^2 = \frac{\sum_{i=1}^n V_i^2}{n-1} - \frac{\langle \delta_{V,i} \rangle^2}{(1 + z_{\text{cl}})^2} , \quad (3.8)$$

where we assume an underlying Gaussian galaxy velocity distribution. In the above equation n is the number of galaxies in the sample and $\langle \delta_{V,i} \rangle$ is the mean measurement error of the galaxy velocities, calculated from $\delta_{V,i} = c \delta_{z,i}$ in which $\delta_{z,i}$ is the uncertainty in each redshift measurement.

Given the random galaxy velocities, their positions and velocity directions are uncorrelated, and we can convert the three-dimensional velocity dispersion and virial radius according to

$$\sigma^2 = 3\sigma_V^2 \quad \text{and} \quad r_{\text{vir}} = \frac{\pi}{2} R_{\text{vir}} , \quad (3.9)$$

$$\text{where} \quad R_{\text{vir}} \equiv 2M^2 \left(\sum_{i \neq j} \frac{m_i m_j}{R_{ij}} \right)^{-1} \quad (3.10)$$

and R_{ij} is the projected separation between the i -th and the j -th galaxy. Applying the above equations to eqs. (3.3) and (3.5) and inserting into eq. (3.1) we find that the virial mass of the galaxy cluster, i.e. the mass within R_{vir} , is

$$M_{\text{vir}} = \frac{3\pi\sigma_V^2 R_{\text{vir}}}{2G}. \quad (3.11)$$

Finally, note that a virial analysis can only be utilised if several tens of galaxies are probed, as otherwise the virial state of the cluster cannot be reliably verified. For example, when applying this method to fewer galaxies it is difficult to account for projection effects, e.g. a merger between two clusters might be overlooked, and the mass estimate will not be reliable.

3.3.2 X-ray

A different method to determine cluster masses comes from measuring the X-ray emission from the hot intracluster plasma that fills the deep gravitational potential wells of galaxy clusters. Assuming this gas is in hydrostatic equilibrium, spherically symmetric, and supported by thermal pressure, the gas density and temperature profiles yield an estimate of the cluster mass (Bahcall & Sarazin 1977; Mathews 1978). This method has several advantages over virial mass estimators, due to the possibility of observationally testing some of the assumptions made.

We know from observations that the larger a galaxy cluster is, the more massive it is, and the higher are its velocity dispersion, X-ray gas temperature, and luminosity. The relation between these parameters can be deduced from theory. The X-ray temperature, T_X , is defined as the thermal energy per gas particle, and should be proportional to the binding energy for a cluster in virial equilibrium,

$$T_X \propto \frac{M_{\text{vir}}}{r_{\text{vir}}}. \quad (3.12)$$

Theoretical considerations of cluster formation find that the virial radius is defined as being the radius within which the average mass density of the cluster is ~ 200 times the critical density of the Universe (see eq. (1.31)). Equation (2.87) yields the virial mass of a cluster, which then leads to

$$T_X \propto \frac{M_{\text{vir}}}{r_{\text{vir}}} \propto r_{\text{vir}}^2 \propto M_{\text{vir}}^{2/3}. \quad (3.13)$$

Observations confirm that a cluster's X-ray temperature provides a very precise measure of its virial mass, and that the $M_{\text{vir}} - T_X$ relation is much tighter than the relation between M_{vir} and the cluster's velocity dispersion.

The total X-ray luminosity that is emitted via bremsstrahlung, L_X , scales as

$$L_X \propto \rho_{\text{gas}}^2 T_X^{1/2} r_{\text{vir}}^3 \propto \rho_{\text{gas}}^2 T_X^{1/2} M_{\text{vir}}, \quad (3.14)$$

where the gas density is given by $\rho_{\text{gas}} \sim M_{\text{gas}} r_{\text{vir}}^{-3} = f_{\text{gas}} M_{\text{vir}} r_{\text{vir}}^{-3}$, in which $f_{\text{gas}} \equiv M_{\text{gas}}/M_{\text{vir}}$ denotes the gas fraction with respect to the total cluster mass. Inserting for eq. (3.13) we then write

$$L_X \propto f_{\text{gas}}^2 M_{\text{vir}}^{4/3}. \quad (3.15)$$

Note that for observations made within a specific energy interval a modified scaling relation must be applied. From observations we know that there is a strong correlation between luminosity and mass, though with a larger scatter than in the mass-temperature relation. The mass of a galaxy cluster is therefore more accurately measured using T_X , rather than L_X . The luminosity in a fixed energy range is, on the other hand, directly observable and hence much easier to determine than measuring the temperature, which requires significantly longer exposure times followed by a fitting procedure to the cluster's X-ray spectrum.

3.4 Structure formation

Cosmic inflation is able to explain the flatness, homogeneity, and isotropy of our observable Universe (see Sect. 1.8.4 for details). Furthermore, it predicts that the primordial Universe had very small perturbations, being the beginning of structure formation in the later Universe. These perturbations have been observed by WMAP as tiny temperature fluctuations in the cosmic microwave background (CMB; see below) radiation. The observed perturbations are also adiabatic, in addition to forming a scale-invariant Gaussian random field, both of which are also predicted by inflation theory.

3.4.1 The early Universe

The Universe was radiation dominated during the early time of its existence, consisting of a mix of various elementary particles which were all interacting with each other. Contributing to the radiation density of the Universe (ρ_r) were the relativistic particles; electrons, positrons, photons, and neutrinos. The non-relativistic particles comprised protons, neutrons, and dark matter. Apart from the dark matter particles, all particle species were in equilibrium by a number of reactions. Since the equilibrium conditions, specified by temperature, continuously change due to expansion, the particles can only maintain equilibrium if the reaction rate (the number of reactions per particle per unit time) is larger than the cosmic expansion rate, $H(t)$. As the reaction rate decreases faster than the expansion rate, particles will at some point no longer be in equilibrium with the other particles and hence decouple from them.

The Big Bang nucleosynthesis (BBN) was initiated roughly 1 second after the Big Bang. Due to the high temperature and plasma density, some of the lightest atomic nuclei were formed from proton and neutron fusion during this period, in particular deuterium (H^2 or D) and helium and lithium isotopes (He^3 and He^4 , Li^6 and Li^7). Being based on strong interactions, the formation of D occurs very efficiently. However, at the time of neutrino decoupling the temperature, T , was not much smaller than the binding energy of D, E_b . Due to the high abundance of photons compared to baryons, a sufficient number of highly energetic photons ($E_\gamma \leq E_b$) existed, which by photo-dissociation instantly destroyed newly formed D. The abundance of D could therefore only become substantial once the temperature had decreased considerably, $k_B T \ll E_b$. About 3 minutes after the Big Bang the D formation rate exceeded the photo-dissociation rate, after which practically all neutrons became bound in D, and He^4 started forming. Due to its high binding energy, He^4 cannot be destroyed by photo-dissociation. All but a small remaining fraction of D quickly transformed into He^4 . The BBN took place during the first few minutes after the

Big Bang, after which the density and temperature of the Universe became too low for nuclear fusion.

The BBN theory predicts that, at around 3 minutes after the Big Bang, the baryonic component of the Universe was composed of $\sim 75\%$ H^1 , $\sim 25\%$ He^4 , and $\sim 0.01\%$ D , with traces of He^3 and Li^7 . These predictions agree with observations of material of low metal content (e.g. metal-poor intergalactic gas), which is thus scarcely affected by nucleosynthesis in stars, and are hence considered as strong evidence for the Big Bang theory. Since no stable nucleus of mass number 5 or 8 exists, no stable, heavier nuclei could form during BBN. Furthermore, the density of He^4 and D depends on the baryon density, Ω_b , in the Universe. This is because a higher Ω_b would lead to a higher baryon-to-photon ratio, which again means that D forms earlier when fewer neutrons have decayed, resulting in a larger fraction of He^4 . Similarly for D , a higher Ω_b leads to a more efficient and complete conversion of D into He^4 , leaving fewer D without a reaction partner, resulting in a lower fraction of D . The relative strength of the $\text{Ly}\alpha$ lines of H (hydrogen) and D can be measured in absorption lines of quasi-stellar objects, yielding a D/H ratio close to the primordial value. Using a Hubble constant of $H_0 \sim 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ renders a baryon density of $\Omega_b \approx 0.04$. Since we know that the matter density of the Universe is $\Omega_m > 0.1$ (Sect. 1.5), this means that most of the matter content in the Universe is non-baryonic dark matter.

After the BBN ended the Universe consisted of photons, electrons, protons, neutrinos, helium nuclei, traces of other light elements, and dark matter. Neutrinos and photons dominated the energy density and hence also the expansion rate. At $z = z_{\text{eq}} \approx 23\,900\Omega_m h^2$, when the energy density of matter equalled that of radiation (Sect. 1.5), pressureless matter started dominating the energy density and the expansion rate. As the Universe continued to cool and expand, free electrons and nuclei could combine into neutral atoms, suppressing the Thompson scattering of photons, resulting in a transparent (i.e. electrically neutral) Universe and the release of the CMB radiation (i.e. free streaming of photons). This process is called recombination or “the last scattering surface”. It occurred at $z_{\text{rec}} \sim 1000$, about 400 000 years after the Big Bang.

The matter in the Universe was almost completely neutral after recombination. However, since we have observed ultra-violet photons from sources at high redshift, $z \sim 6$, we know that the intergalactic medium at lower redshift must be highly ionised, otherwise these photons would have been absorbed by photo-ionisation of neutral hydrogen. At some point between $z \sim 1000$ and $z \sim 6$ the intergalactic medium must therefore have been re-ionised, most likely by the first active galactic nuclei or the first stars.

3.4.2 Linear perturbation growth

During the 20th century it was discovered that the major matter component of the Universe was not atoms, but an unknown form of matter called dark matter (see also Sect. 1.8.1). Comparing the BBN predictions of the baryon density in the Universe today with the current total matter density tells us that this “new” type of matter is not composed of baryons. In addition we know that it does not emit or absorb radiation, however it does react through the force of gravitation. Its components may be of a type that weakly interacts (like neutrinos), although they cannot consist entirely of the three known kinds of neutrinos. This comes from the fact that neutrinos are relativistic parti-

cles, not gravitationally bound in the potential well of a density concentration. Such hot dark matter yields structure formation which cannot explain the large-scale structure we observe in the Universe today. Since it mostly interacts through gravitation, cold dark matter plays a key role in structure formation. Now the Jeans instability, which denotes the situation when the internal gas pressure of a matter-filled region is not strong enough to withstand a gravitational collapse, is not opposed by radiation or any other force. Dark matter therefore collapses into dark matter haloes much sooner than ordinary matter, for which the process is delayed by pressure forces. Dark matter is hence able to explain the matter structure we see today – structure formation would occur much later if dark matter did not exist, and the Universe would not contain the majority of the structure we see today.

Dark matter perturbations evolve independently as a function of scale, comprising larger and larger perturbations as the Universe expands. We know from relativistic perturbation theory that density fluctuations grow, due to gravitational clustering, as long as their comoving scale is larger than the comoving horizon scale ($r_{\text{H,com}}$; see Sect. 1.8.3), which is $\propto a^2$ in the radiation-dominated Universe, and $\propto a$ during the matter-dominated phase. However, if the horizon scale becomes larger than the perturbations during the period when the energy density is dominated by radiation, efficient perturbation growth is prevented due to the fast radiation-driven expansion rate of the Universe, and can only continue once the Universe becomes matter-dominated. Since light can only cross regions that are smaller than the horizon scale, the suppression of growth due to radiation is therefore restricted to scales smaller than the horizon.

The evolution of baryonic density fluctuations differ from that of dark matter. Despite the matter-dominance, the baryon density remains smaller than that of photons until after the beginning of recombination. Baryons are hindered from falling into the potential wells of dark matter due to their strong interaction with photons. They are hence subject to radiation pressure, leading to an initially much smoother density distribution than that of dark matter. After recombination this interaction comes to an end, and baryons can then fall into the potential wells of dark matter. At a later stage the distribution of baryons will therefore closely resemble that of dark matter.

3.4.3 Nonlinear growth of structure

As the perturbations grow, dark matter regions become too dense (\gg the mean density of the Universe) to be treated within the framework of linear perturbation theory. Furthermore, as the dark matter particles start forming gravitationally bound systems, dark matter can no longer be treated as a pressureless fluid. The linear and fluid approximations utilised until now become invalid, and the Newtonian theory of gravitation (Sect. 1.2) must be applied. Agreeing largely with observations, N-body simulations are generally the best tool for understanding such dynamics. Forming galaxy groups, clusters, and superclusters, these simulations show that matter concentrates in a web-like structure of filaments and haloes, whilst most of the Universe consists of voids.

Finally, the last step of evolution happens as baryons are compressed in the centres of galaxy haloes, forming stars and galaxies. Although the formation of dense haloes is greatly accelerated by dark matter, since it does not have radiation pressure, dark matter cannot form smaller structures as it cannot dissipate angular momentum. Dense objects

Chapter 3. Clusters of galaxies

are therefore created from the collapse of ordinary baryonic matter which can dissipate angular momentum through radiative cooling. Including physics from several different fields, simulations of these processes are very difficult to perform.

Chapter 4

Data reduction

The essence of observational astronomy lies in detecting electromagnetic radiation. Technology has greatly enhanced our possibilities to collect and record as much light as possible, starting with Galileo's discovery of the telescope four centuries ago and later the invention of photography. The two combined allowed for the detection of very faint objects, inspiring further technological improvements. The invention of charge-coupled devices (CCDs) in the 1970s once more revolutionised astronomy, capturing $\sim 80\%$ of the incoming light as opposed to $\sim 2\%$ for photographic plates, in addition to having a linear response. The introduction of CCDs also significantly improved spectroscopic studies, allowing for observations of much fainter sources than previously possible. The technological advance in spectrographs in the later years has in addition enabled large spectroscopic surveys, rendering precise redshifts of thousands of objects. The development of photometric redshift methods finally allowed for a large-scale, three-dimensional mapping of the Universe, opening up many new research areas in observational cosmology.

In addition to collecting light from astronomical sources, the CCD detector is also subject to technical issues regarding telescope and camera which must be corrected for. This calibration is called image reduction, and is presented in this chapter. It specifically serves to remove the instrumental signature in addition to improving the signal-to-noise (S/N) in the data, and must be applied to all astronomical CCD images before scientific information can be extracted. The process of redshift determination will be explained towards the end of the chapter.

Howell (2006) gives a detailed and technical summary of the topic, whilst Schneider (2006b) provides insight into the subject from a weak lensing point-of-view. Further details regarding redshift determination can be found in Schneider (2006c).

4.1 Pre-Processing

Glossary A few terms often used in observational astronomy are listed below together with an explanation.

FITS Flexible Image Transport System; a digital file format specifically designed for scientific data, to store, transmit, and manipulate scientific images.

Exposure A single image, consisting of one or more CCD frames.

Chapter 4. Data reduction

Integration Taking an exposure – integration time is the same as exposure time.

Pixel A CCD comprises an array of picture elements called pixels, made from silicon. Incoming photons in the [380 – 1000 nm] wavelength range are absorbed by the silicon, which correspondingly releases an electron. The released electron is stored in a potential well within the pixel until the end of the exposure, when the collected charges in the CCD are shifted along each column and read out row by row.

Saturation A pixel reaches saturation either when

- * it has been exposed to too much light and overflows (so-called blooming; excess photoelectrons bleed into neighbouring pixels), or
- * the pixel has stored more photoelectrons than can be represented in output digital numbers (ADU; analog-to-digital units) by the A/D (analog-to-digital) converter.

Dithering The telescope is offset between exposures to fill gaps between chips (see below), and to avoid that light from objects of interest fall onto bad pixels in all exposures (and hence get masked out, see Sect. 4.2).

Multi-chip mosaic camera A camera containing more than one CCD (each CCD is then called chip). Note that for a multi-chip detector an individual set of calibration frames must be obtained and applied to each chip separately, i.e. the chips are calibrated independently until Sect. 4.2.3.

Stacking Different exposures from the same chip are combined on a pixel-to-pixel basis; each pixel's ADU values from the different frames are averaged using the mean (or median if specified), yielding that pixel's output value in the combined frame.

Normalisation The pixel values of a frame are rescaled with the mode of the image (the normalisation factor), resulting in an output image with mode equal 1. The mode of an image is defined as the most common pixel ADU value; a histogram representing all pixel values is smoothed to eliminate noise, after which the most common value is chosen.

Note that the calibration steps are presented in their sequential order in the next two sections.

4.1.1 Bias

All pixels in a raw frame contain a positive offset value called the bias level. Due to read noise (see below) this level is set in order to avoid negative values in the output image (mainly a problem in short exposures), such that no bit in the A/D converter must be used to store information about the sign but all bits can be used with respect to the number of photons in each pixel.

The two-dimensional variations in the bias level are measured by reading out a frame with zero exposure time, leaving the camera shutter closed. To eliminate cosmic rays, read noise variations, and random fluctuations, ideally 10 (or more) individual frames are median stacked using outlier rejection. Since the bias level is an additive effect the master bias frame, e.g. Fig. 4.1a, is subtracted from all frames including calibration frames.

Although infrequent, the bias level in some CCDs varies during the night. This problem is corrected for by adding a number of pseudo-columns and/or rows to each frame. These overscan regions are not physical pixels but generated by the output CCD electronics. Upon subtraction of the master bias, the mean of the overscan regions must be subtracted from each respective frame (including the bias frames before creating the master bias). If no bias frames were taken, the overscan regions can be used to remove the bias. Note, however, that this method is only sufficient if it is clear that the CCD does not suffer from a two-dimensional bias pattern. The data used in Chap. 5 has a very smooth bias, hence this method could successfully be applied without biasing our analysis (Sect. 5.1.2).

Using the above procedures the bias level is generally 100% corrected for.

Read noise The read noise of a CCD describes the uncertainty in the readout process. It consists of two components; the conversion from an analog signal to a digital number, and the introduction of spurious electrons by the electronics causing random fluctuations in the output. The read noise denotes the 1σ uncertainty and can be estimated from the rms (root mean square) measured in several single bias frames.

4.1.2 Dark current

CCDs are subject to thermal noise, also called dark current noise, since their temperature is much larger than 0K. Due to this effect electrons are freed from atoms in the silicon of a CCD, and get trapped in pixels' potential wells as if released by an incoming photon. Most CCDs are cooled to about -120°C to mitigate this effect (however defect pixels can show significantly higher dark current than good pixels).

Though adding to the overall noise in the image, the dark current itself is easily corrected for. A dark current frame is created by integrating over the same exposure time as in the images it will be used to correct for, while leaving the shutter closed to ensure that no outside photons hit the detector. Thus only the dark current is registered during the exposure. A master dark frame is created by stacking 10 or more single frames, which is then subtracted from all science and calibration frames. Note that this frame also contains the CCD's bias level, hence the subtraction of a master bias can be omitted if this method is applied. If, however, the exposure time of the dark frames does not equal that of all the science frames, the single dark frames must be bias corrected upon stacking. This master dark is then normalised, after which it is multiplied with the exposure time of the science frame to be corrected upon subtraction.

The dark current subtraction is exact up to its Poissonian noise.

4.1.3 Flatfield

Each pixel within a CCD has a slightly different quantum efficiency (QE; a measure of how efficiently the CCD detects and converts photons to ADU). Each pixel's relative response to incoming photons must therefore be flattened, which is done through the division of a master flatfield. The master flatfield is created by first normalising all single flatfield frames and then stacking them. An ideal flatfield is one in which all pixels are uniformly lit by a light source comprising the same spectral energy distribution as that of the science frames, yielding a spectrally and spatially flat image.

In ground-based observations there are two standard methods used to obtain a flatfield. Dome flats are created by uniformly illuminating either the inside of the telescope dome or a screen mounted there, and taking short exposures to avoid saturation of the CCD. Sky flats are often the preferred method, taken during twilight when the sky is smoothly lit. Note that the sky must be free from clouds in order to obtain uniformly lit flatfields free from structure. The telescope is pointed to a so-called blank field (containing no bright stars leaving large extended haloes) and dithered between exposures. Visible stars are automatically excluded from the master flatfield frame through a median stacking.

Obtaining flatfields from space-based observations is a more complicated process. One option is to use flats taken prior to launch – however this is not an optimal solution as the CCD response often changes over time. A better option is to use defocussed or scanned observations of the bright Earth or Moon, or internal calibration lamps.

Because the pixels' response to incoming light is wavelength-dependent, one master flatfield must be created for each filter (Sect. 4.3.2) used in the science observations, see Fig. 4.1b. This is also important due to the fact that dust settles on the filters in different places. Filters are usually also slightly damaged from cleaning or coating processes, which is different from filter to filter. Each filter therefore has an imprint that must be taken into account. Every master flatfield should be created from at least 5 single frames taken in the corresponding filter, with an exposure time resulting in ADU counts of about half the amount leading to saturation.

In addition to correcting for the individual pixel's response to light, a flatfield will also correct for large-scale throughput variations. This effect arises mostly due to vignetting of improperly sized filters or lenses in the light beam. Figure 4.1c shows a science image which has been bias and flatfield corrected.

Superflat In addition to flatfielding, a so-called superflat can be computed by normalising all science frames (taken in the same filter) to the same brightness level (not necessarily 1) and median stacking them. The superflat has the advantage of being taken during night-time, capturing the spectral energy distribution in the science frames. Utilising a median stacking again ensures that objects are filtered out (given the images were dithered), however, for the same reason note that science observations of extended sources cannot be used to compute a superflat. The superflat is heavily smoothed and normalised (to 1) to provide an illumination correction comprising final corrections not contained in the master flatfield. All science exposures are then divided by this illumination correction, an example of which is shown in Fig. 4.1d.

The flatfielding process accounts for 90 – 100% of the sensitivity variations within and between chips, depending on the camera in question. If an illumination correction is included in the process the intra-chip variations are generally 100% corrected for. The inaccuracy of the inter-chip corrections usually increase the larger the field-of-view, see also Sect. 4.2.3.

An unsmoothed superflat can be used as an alternative to dome or sky flats if none have been taken. However, in this case the science frames are subject to strict requirements. The exposure time of each frame must be long enough to reach the needed S/N ratio, and the dithering must be larger than the diameter of the observed objects to ensure they are filtered out during the superflat creation. The accuracy of this procedure is discussed in detail in Sect. 5.1.2.

4.1.4 Fringing

When observing monochromatic light, a pattern called fringing may appear in the CCD image. Fringes are created from interference between photons reflected within the CCD or between long-wavelength light rays passing through the detector and being reflected back into it. Fringing usually occurs in red narrow- and broadband filters (Sect. 4.3.2), like the R , I , or z band, where the night-sky spectrum is dominated by atmospheric OH emission lines. These forbidden transitions are powered during day-time by ultraviolet sunlight and have decay times of many hours. They are highly variable in time and strength.

Since the reflected sunlight in the twilight sky has a continuum spectrum, the narrow emission lines dominating the dark night sky are only very weakly present in a sky flat. A superflat must therefore be used for fringe modelling. Subtracting a strongly smoothed superflat from the original version leaves only the additive effect of fringing, see Fig. 4.1e. In order for the amplitude of the fringes to be compatible with the science image in question, the fringe model is then rescaled to the sky background level (Sect. 4.2.1) of the corresponding science frame upon subtraction. Figure 4.1c shows a science frame before fringe correction, whereas in Fig. 4.1f the fringes have been cleanly subtracted. Note that if the atmospheric conditions at night are unstable, several fringe models must be created from smaller data sets taken during different parts of the night with similar conditions, and used to correct the corresponding science frames. For example, a separate fringe model can be generated for every hour of the night, each created from and used to correct the exposures taken within that hour.

In order to adequately capture the fringes and reduce the noise in the background sky it is often necessary to obtain long exposures of blank fields several times throughout the night. This was done during our observations of the galaxy clusters analysed in Chap. 5, see the beginning of Sect. 5.1.2. During unstable conditions these blank field observations must be performed on a regular basis.

Following the procedure presented above, the fringe subtraction is 95 – 100% accurate.

4.2 Image combination

4.2.1 Sky subtraction

In addition to receiving light from the target of interest, the CCD also collects radiation from the background sky (skyglow), as well as from undetected objects, moonlight, and potential light pollution from nearby civilisation. This background level must be subtracted from each image in order to solely measure the flux of the source. In the case of a multi-chip camera, the procedure must be performed before image coaddition (Sect. 4.2.4) to avoid discrete jumps in the background level across chip borders.

To create a background sky model, all objects in each science frame are removed and the remaining image smoothed with a large kernel width, usually of about 100 – 500 pixels. This sky model is then subtracted from the pre-processed science frames.

4.2.2 Bad pixel mask

Bad pixels and columns, hot pixels, and cosmic rays appear in the science images, creating problems for the astrometric calibration (see below) and the analysis process, and must

thus be flagged. The pixels are assigned some value recognised by the processing software, allowing them to be neglected in further operations with the images. An image in which bad pixels have been flagged (usually set to zero) and good pixels have a value of 1 is called a mask frame, see Fig. 4.1g.

A global mask image is usually obtained from the master flatfield and/or master dark¹, in which the bad pixels can be easily recognised and flagged since they have ADU counts far above or below the mean pixel value. In addition to the global mask, which is applied to all science frames, individual mask frames are often also created. The individual masks are usually made manually, and contain unwanted regions that change from image to image. Examples are saturated stars or satellite tracks. Hot pixels and cosmic rays are detected automatically in individual science frames and added to the individual masks.

Weight image A mask file that has been created using the master flatfield as input image and in which the good pixels have not been set to 1 is called a weight image. It encodes how much information one pixel carries with respect to other pixels, and thus optimises the overall S/N of the final coadded image (eq. (4.2)). By rescaling the individual weight images using the exposure time of the corresponding science frame (eq. (4.4)) and coadding them using the same astrometric solution as is used for the science images (Sect. 4.2.4), a coadded weight image is created. This master weight contains information about the effective total exposure time of each pixel, and hence serves as a noise map from which the S/N of each object can be calculated. The individual weight images can also be included in the process of photometric calibration (Sect. 4.3.3).

4.2.3 Astrometry and relative photometry

During coaddition (see below) all the science frames are coadded to increase the S/N with respect to the individual images. The noise in the master bias and master flatfield (due to the finite number of frames stacked) leaves an identical imprint on the science images during pre-processing. Dithering the telescope between exposures hence ensures that this noise averages out during coaddition, increasing $\propto \sqrt{N}$. If no dithering is applied the noise will increase $\propto N$ as does the flux of the objects. Light from the same object therefore hits different places on the CCD in different exposures, thus the coaddition cannot simply be done on a pixel-to-pixel basis. In addition, influences from the telescope and camera optics like telescope distortion, thermal expansion, and mechanical strain in addition to atmospheric refraction lead to field distortions that must be corrected for – this effect is particularly visible in multi-chip images with a wide field-of-view. An accurate non-linear mapping from pixel coordinates to sky coordinates must therefore be calculated. These distortions usually change slowly over time and must hence be calculated from the data themselves.

On a chip-to-chip basis, objects are detected in the science frames, after which they are matched to objects in a reference catalogue containing accurate sky coordinate information. Objects inside the masks created in the previous section are ignored during this process. An astrometric solution is then computed – first a linear mapping between reference points in each exposure and a reference sky coordinate is calculated (to account

¹For a multi-chip detector one global mask is made for each CCD.

for dithering and rotation), then two-dimensional polynomials are determined in addition to correct for non-linear distortions.

Once the astrometric mapping is known, object fluxes in different images must be compared and adjusted. This relative photometry is necessary as the atmospheric transparency (including airmass; an object's optical path length through the Earth's atmosphere – the airmass in an observation is determined by the pointing angle of the telescope) gradually changes during the night. If I_i denotes the part of an input pixel contributing to an output pixel (I_{out}) in the coadded image, the relative photometry is done by scaling I_i with

$$f_i = \frac{10^{-0.4 ZP_i}}{t_i} \quad (4.1)$$

(Erben et al. 2005), which also includes a normalisation to 1 second. The relative photometric zero-point is denoted ZP_i and the exposure time t_i .

It is assumed that all chips have the same zero-point, i.e. that relative photometry between chips has been corrected for during flatfielding and illumination correction. This is generally a reasonable assumption, however zero-point variations between chips become more critical the larger the field-of-view, up to $\sim 10\% \approx 0.1$ magnitude (Sect. 4.3.1). In order to correct for these variations there are two options. If the relative zero-point between chips has already been measured these corrections can be applied. If such measurements have not been made, the same photometric standard star must be observed in all chips, from which inter-chip photometry can be performed (see Sect. 4.3.3 for details). Note that for multi-chip detectors, zero-point variations between chips are usually monitored and corrected for by observatory staff upon data release.

4.2.4 Coaddition

Finally, the coaddition is made by first applying the astrometric solution to each corresponding science and weight image. In order to ensure that each object falls onto the same pixels, this procedure includes a resampling of each input frame onto a new output frame. These newly resampled frames are then stacked using a weighted mean, taking the individual weight images into account to further increase the S/N , according to

$$I_{\text{out}} = \frac{\sum_{i=1}^N I_i f_i W_i w_i}{\sum_{i=1}^N W_i w_i} \quad (4.2)$$

(Erben et al. 2005). Here N is the number of exposures in the stack, W_i is the input value corresponding to I_i in the weight image, and w_i is a weight,

$$w_i = \frac{1}{\sigma_{\text{sky},i}^2 J_i^2}, \quad (4.3)$$

where $\sigma_{\text{sky},i}$ represents the noise in the sky background in each exposure. An example of a coadded science frame is shown in Fig. 4.1i.

The weight images are stacked following

$$W_{\text{out}} = \sum_{i=1}^N W_i w_i, \quad (4.4)$$

see Fig. 4.1h. Here W_{out} denotes the value corresponding to I_{out} in the coadded weight image.

Object detection The detection of objects in a FITS image can be done using two different methods. The first defines a S/N threshold and counts connected pixels above this threshold, including the object in the catalogue if the number of connected pixels exceeds a chosen lower limit. The other method allows for a range of filters of different radii to smooth the image – the detections with the highest significance are then defined as objects. The weight images are taken into account during object detection as they provide the noise level at a given pixel position. Masked out regions are ignored, thus no objects are detected here.

4.3 Photometry

Photometry is defined by Howell (2006) as “the amount of temporal nature of the flux emitted by an object as a function of wavelength”. A photometric measurement of an object renders its brightness measured in magnitude or luminosity (see below). Photometric studies may yield the absolute photometry of objects, i.e. their absolute flux, and can also be used to measure differential photometry of objects, i.e. their change in brightness over time, yielding the objects’ light curve.

4.3.1 Magnitudes

The magnitude of an object denotes its apparent brightness, i.e. the amount of electromagnetic energy an object radiates per unit area per unit time as measured on Earth. Magnitude is more precisely denoted apparent magnitude, m , to distinguish it from the distance-independent absolute magnitude, M , which is defined as the apparent magnitude an object would have if it was located 10 pc away from us,

$$m - M = 5 \log_{10} \left(\frac{D_{\text{lum}}}{10 \text{ pc}} \right) . \quad (4.5)$$

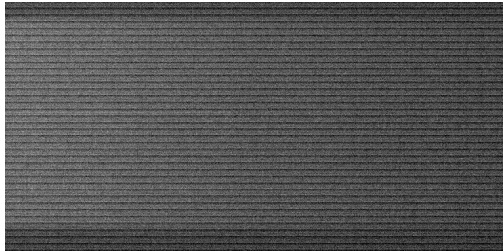
The luminosity of an object represents the amount of energy emitted by the object per unit of time. It is distance-independent and can be calculated directly from an object’s absolute magnitude as

$$L = 10^{-0.4(M - M_{\odot})} L_{\odot} \quad (4.6)$$

(see also Sect. 7.3.1), where M_{\odot} (L_{\odot}) represents the absolute magnitude (luminosity) of the Sun. Luminosity is usually given for a specific filter – if none is specified it represents the bolometric luminosity, i.e. the total energy radiated by an object at all wavelengths. Magnitude is measured on an inverse logarithmic scale, in which the fainter the object the higher its magnitude is, whereas luminosity is measured on a linear scale where fainter objects have lower luminosity.

The observed light distribution of an object can be described as a convolution between its light distribution outside the atmosphere, $I^{\text{intr}}(\boldsymbol{\theta})$, and the point-spread function (PSF) of the imaging system, P ,

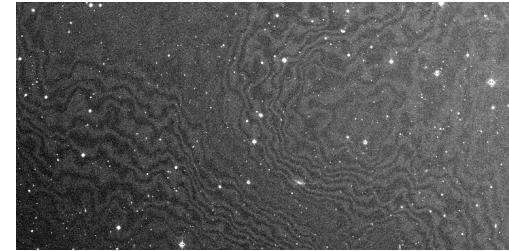
$$I^{\text{obs}}(\boldsymbol{\theta}) = \int d^2\theta' I^{\text{intr}}(\boldsymbol{\theta}') P(\boldsymbol{\theta} - \boldsymbol{\theta}') , \quad (4.7)$$



a: Bias; the values range from -1 to $+2$ ADU.



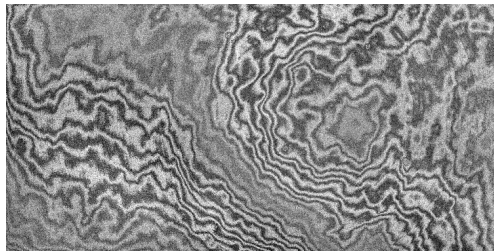
b: Flatfield; the counts drop by $\sim 25\%$ from the upper right to the lower left corner (due to vignetting of the telescope optics).



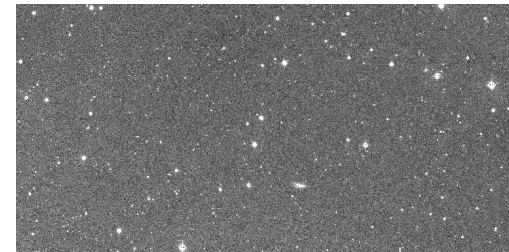
c: Bias and flatfield corrected image; residual background variations (fringing + flatfield) of $\sim 8\%$ can still be seen.



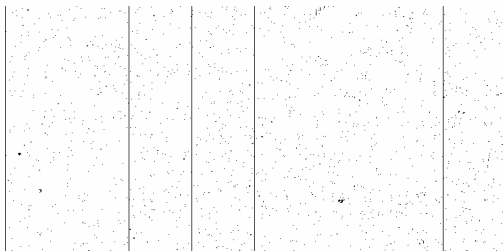
d: Illumination correction; the levels in this image vary by $\sim 7\%$ (300 ADU).



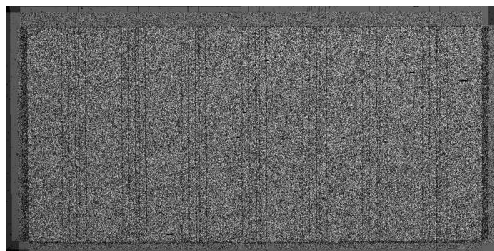
e: Fringe model; the fringing amplitude is $\sim 2.5\%$ of the background level.



f: Fully pre-processed image; the background is flat to within 1%.



g: Bad pixel mask; the black regions have a value of 0 ADU, the white 1 ADU.



h: Coadded weight image.



i: Coadded science image.

Figure 4.1: Separate frames showing the different image reduction and calibration steps, see the text for details.

where $\theta = (\theta_1, \theta_2)$ (see also eqs. (2.1)). Converting into (x, y) detector coordinates we have

$$I^{\text{obs}}(x, y) = \iint dx' dy' I^{\text{intr}}(x', y') P(x - x', y - y'). \quad (4.8)$$

Furthermore, the instrumental flux of an object can be defined by summing up the total brightness collected by all pixels covered by the object,

$$S^{\text{instr}} = \sum I^{\text{obs}}(x, y). \quad (4.9)$$

The difference between the apparent magnitudes, $m_{1/2}$, of two separate sources is defined in terms of their observed fluxes, $S_{1/2}$, as

$$\Delta m = m_1 - m_2 = -2.5 \log_{10} \left(\frac{S_1}{S_2} \right). \quad (4.10)$$

4.3.2 Commonly used filters

A filter, also called a bandpass/band, is transparent to light of certain wavelengths, while blocking or reducing light outside the selected bandpass. The most common broadband filters used in optical astronomy are the U, B, V, R, I Johnson-Bessel filters shown in Fig. 4.2a and the u', g', r', i', z' SDSS (Sloan Digital Sky Survey) filters in Fig. 4.2b, where the z band is approaching the near-IR part of the spectrum. Those used in near-IR observations are typically the Y, J, H, K UKIDSS (UKIRT² Infrared Deep Sky Survey) filters shown in Fig. 4.2c. In addition we have medium- and narrowband filters, in which only a narrower band of the spectrum is transmitted. These filters require long exposure times in order to obtain a sufficiently high S/N .

Photometric redshift estimates utilise observations made of the same objects in different filters to estimate their redshifts. Although broadband observations might be sufficient for a large fraction of the objects, adding mediumband filters will increase their redshift accuracy in addition to decreasing the number of catastrophic outliers (objects whose redshift estimates are completely wrong). In Sect. 4.4 we explore the topic of redshift determination.

Weak lensing analyses require deep observations of faint higher-redshift galaxies in excellent seeing conditions. Since the seeing decreases slowly with increasing wavelength, r and i bands are preferred for weak lensing observations. Furthermore, objects show less substructure in these bands as compared to blue filters, increasing the robustness of the shape determination. Observations made in z band are about 1 – 1.5 magnitudes less deep due to reduced detector QE and increased sky background. Because of the high sky background in the IR regime, observations here are usually not deep enough to apply weak lensing – however, near-IR observations at large ($\gtrsim 8$ -meter) telescopes using IR detectors are currently being tested for weak-lensing applications (T. Schrabback & M. Schirmer, priv. comm.; King et al. 2002).

²United Kingdom Infrared Telescope

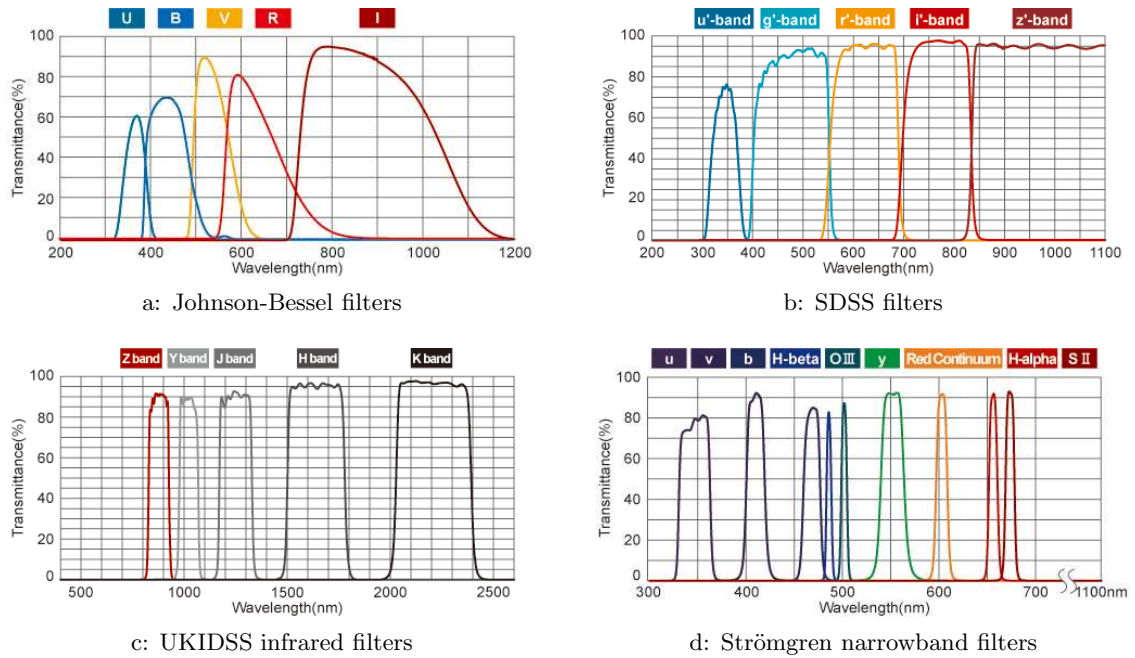


Figure 4.2: Transmittance curves for optical, infrared, and narrowband filters. The transmittance indicated by the y -axis is defined as the percentage of the light falling onto the filter that actually passes through it. (Figure credit: http://www.asahi-spectra.com/opticalfilters/astronomical_filter.html)

4.3.3 Photometric calibration

The simplest method for measuring the flux of an object is aperture photometry. This technique simply sums up the observed (background subtracted) brightness counts within an aperture centred on the source, as described in eq. (4.9). This aperture makes no assumption regarding the shape of the PSF, and can take on any shape needed. Aperture photometry is very useful given its computationally simple form, and is often used together with a set of different sized and/or shaped apertures, where the best-fit aperture is chosen individually for each object according to a set of selection criteria (this method was used in “Astrometric calibration” under Sect. 5.1.2). The optimum aperture is commonly found to yield the highest S/N of the object. Magnitudes measured using aperture photometry are often called aperture magnitudes, especially if the applied aperture is smaller than the size of the object, as is the case in for example photometric redshift determination (Sect. 4.4.1). Note that aperture photometry is not applicable to blended sources, for which the light distribution should rather be fitted to avoid including light from nearby objects.

Accurate absolute photometry requires additional observations of photometric standard stars. These are stars whose magnitudes have been carefully measured in several filters and are publicly available, e.g. Landolt (1983, 1992; Landolt magnitude zeropoints) and Stetson (2000). Offsets are then found between these stars’ observed magnitudes and those found in the photometric standard star catalogue(s), accounting for atmospheric extinction and small differences in the bandpasses. Applying this photometric calibration

to the science frames hence provides a method for calibrating all object magnitudes. Note that in order to perform accurate photometry, both science and standard star observations must be performed in photometric conditions, i.e. a clear, cloudless sky, where the atmospheric extinction is a function of airmass only.

4.4 Redshift determination

Due to the expansion of the Universe, (most) galaxies' spectral lines are redshifted towards longer wavelengths (Sect. 1.1). Redshift can thus be used for distance determination. Two methods exist for this purpose – spectroscopic and photometric redshift determination.

4.4.1 Photometric redshifts

Photometric redshift (photo-z) determination of galaxies is based on multicolour photometry. A galaxy's broadband energy distribution is a superposition of stellar and nebular radiation, showing characteristic features dependent on redshift. The redshift of a galaxy can therefore be determined from observations in different filters using the following procedure.

Photo-z determination is based on a set of standard galaxy spectra (templates) collected from observed galaxies or calculated based on simulations of the evolution of stellar populations. Each template spectrum is then redshifted, from which the expected galaxy colour for any given spectrum and redshift can be computed for all filters. These colours are then compared to the observed galaxy's colours, from which a best-fit redshift, as well as galaxy type, is inferred.

4.4.2 Spectroscopic redshifts

Spectroscopic observations encompass the dispersion of electromagnetic radiation into its component energies (e.g. visible white light will be dispersed into different colours). The emitted energy distribution from an object is collected in wavelength bins of size $\Delta\lambda$ – this spectral resolution is determined according to the observer's scientific objective.

Since the emission and absorption spectra of atoms are distinctive and well known, spectroscopic observations provide precise redshift information about the object(s) in question. For elliptical galaxies with old stellar populations the absorption lines Ca H (3934 Å) (commonly denoted Ca H λ 3934) and K λ 3969, or Mg λ 5179 and Na λ 5893, are usually utilised. Star-forming galaxies exhibit distinct emission lines such as [OII] λ 3727, H β λ 4865, [OIII] λ 5007, and H α λ 6564, where brackets denote forbidden transitions.

The working principle of a spectrograph is described in the following. After being collected by the telescope, light from one object is focussed on a slit³. The width of the slit both sets the spectral resolution and eliminates unnecessary background light. Next, the light is dispersed by a grating, prism, or a combination of the two (grism), before the spectrum is focussed once more onto a CCD. As for CCD imaging, a CCD spectrum must be reduced as described in Sects. 4.1 and 4.2. However, since spectroscopic observations are very sensitive to mechanical strain in the telescope as well as airmass, wavelength

³In multi-object spectroscopy (several short slits covering different objects whose spectra are all captured in the same exposure) the following procedure applies to all objects separately.

calibration using the exact same setup must be taken immediately before and after each scientific observation in order to correct for small wavelength shifts.

4.4.3 Spectroscopic vs. photometric redshift determination

Both redshift determination methods have advantages and disadvantages. Choosing one method depends on the scientific goal of the project, as well as the access to observing time and archival data.

Spectro-z observations are in general much more observationally expensive than photo-z observations, since obtaining spectra of individual galaxies is more time-consuming and complex than multicolour photometry. Photo-zs can in addition be extended to faint magnitudes – objects too faint to be captured in a dispersed spectrum. Photo-zs may also be estimated for far more galaxies at the same time than that of spectro-zs, given that in principle all objects in the image are subject to redshift calculations. On the other hand, newer spectrographs in which the spectra are fibre-fed onto the CCD also renders the possibility of obtaining as many as a few hundred objects in the same exposure.

Spectro-zs are much more accurate than photo-zs. This is due to both the unambiguity of a spectrum, as well as the assumptions made by the photo-z method regarding the nature of the object’s spectrum in addition to relying on filters that are sensitive to a range of wavelengths. For early-type galaxies forming a well-defined colour-magnitude sequence at any redshift (Sect. 3.1), the photo-z method works very well due to the galaxies’ tight relation between redshift and multicolour information. However, if the distinct, so-called Balmer break at $\lambda = 4000 \text{ \AA}$ is not located between two of the employed filters, this relation is no longer unambiguous. The spectral energy distribution of other galaxy types include larger variations, increasing the uncertainty of the photo-z measurements.

The number of filters and their wavelength coverage is crucial for photo-z determination. Given an insufficient number of filters the redshift estimate can be completely wrong. To avoid such catastrophic outliers the spectral energy distribution over a wide range in wavelengths must be probed by an adequate number of filters ($\gtrsim 5$ photometric bands). Combining optical and near-IR filters is recommended, and even critical in photo-z determination of high-redshift galaxies. In the latter case, for $z \gtrsim 1$ the 4000 \AA break is redshifted beyond 8000 \AA , complicating a photo-z calculation from optical bands only.

Weak lensing density profiles and mass reconstructions of the galaxy clusters Abell 1351 and Abell 1995

Clusters of galaxies comprise the most massive gravitationally bound and relaxed structures in the Universe. They thus represent a vital tool in our search for a deeper understanding of the properties of dark matter (Chap. 3). Gravitational lensing studies do not depend on the nature or the dynamical state of the gravitating matter, providing a powerful method to measure the mass of galaxy clusters independent of their baryonic content (Chap. 2).

In this chapter we study the overall mass distribution of the intermediate-redshift galaxy clusters Abell 1351 and Abell 1995 using weak gravitational lensing. These clusters have a very different mass structure and dynamical state, and in this respect represent the two extremes of a larger sample of 38 highly X-ray luminous clusters of similar size and redshift previously studied by Dahle et al. (2002). Re-observing the clusters with the wide-field camera CFH12K at the Canada-France-Hawaii telescope (CFHT) provided a larger field of view than employed by Dahle et al. (2002), allowing us to map the clusters to larger radii than previously possible.

The KSB+ method (Sect. 5.2) is used to recover the shear values of faint background galaxies in the images. Using a finite-field reconstruction technique (Sect. 5.3), we derive two-dimensional mass maps, visualising the surface mass distributions of Abell 1351 and Abell 1995. We also apply aperture mass statistics (Sect. 5.3.1) to our data, comparing the results to confirm mass peak detections. Finally, by fitting predicted shear values from theoretical models to the shapes of the lensed galaxies we estimate the cluster masses (Sect. 5.4).

This chapter is largely based on Holhjem, Schirmer, & Dahle (2009). Throughout this thesis we assume a Λ CDM cosmology (Sect. 1.8), with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $h_{70} = H_0/(70 \text{ km s}^{-1} \text{ Mpc}^{-1})$. All errors represent the 68% confidence level except where otherwise specified.

5.1 Observations and data reduction

5.1.1 Data acquisition

The galaxy clusters Abell 1351 and Abell 1995 are centred at the positions $11^{\text{h}}42^{\text{m}}30^{\text{s}}.7 + 58^{\circ}32'21''$ and $14^{\text{h}}52^{\text{m}}50^{\text{s}}.4 + 58^{\circ}02'48''$, respectively. They were observed with the CFHT on the 4 nights of 7-11 May, 2000, using the wide-field CCD mosaic camera CFH12K. This camera covers a field of $12\text{k} \times 8\text{k}$ pixels in total, representing an area of $42'.2 \times 28'.1$ on the sky. The pixel scale is $0''.206$ when mounted at the CFHT prime focus. A total exposure time of 5400s was obtained for both clusters in I band. However, due to seeing $> 1''$, 3 exposures were rejected from the Abell 1351 data set, resulting in 3600s for this cluster. This corresponds to a 5σ limiting magnitude of $I = 25.2 \pm 0.1$ for point sources in both pointings. The seeing in the final coadded images equals $0''.95$ and $1''.15$ for Abell 1351 and Abell 1995, respectively. The number density of the lensed background galaxies is 16 arcmin^{-2} for both clusters. The ellipticity dispersion (after point-spread-function correction, PSF; Sect. 5.2) is $\sigma_g = 0.43$ and $\sigma_g = 0.51$ (eq. (2.46)) for Abell 1351 and Abell 1995, respectively. The larger dispersion for Abell 1995 is explained by the 20% higher image seeing, which enlarges the PSF correction factors and their uncertainties.

In addition we made use of the V -band data obtained by Dahle et al. (2002) to verify neighbouring peaks present in our two-dimensional mass maps described in Sect. 5.3. These data were obtained at the 2.24m University of Hawaii telescope using the UH8K mosaic camera, covering an area of $4\text{k} \times 4\text{k}$ pixels (rebinned 2×2) mapping $18'.8 \times 18'.8$ of the sky. Each image has a total exposure time of 12 600s, resulting in a depth comparable to our I -band data, with 5σ limiting magnitudes of $V = 25.9 \pm 0.1$ and $V = 25.8 \pm 0.1$ for Abell 1351 and Abell 1995, respectively. Further details about the reduction process and coaddition of the V -band data can be found in Dahle et al. (2002).

Data processing and analysis are carried out using mainly the IMCAT software package¹, Kaiser's July 2005 version for Macintosh. IMCAT is a tool specially designed for weak lensing purposes, and is optimised for shape measurements of faint galaxies. It processes both FITS files (Sect. 4.1) and object catalogues.

As Abell 1351 and Abell 1995 are both located at redshift $z = 0.32$, they have a similar correspondence between physical and angular scale, given as $1 \text{ Mpc} = 215''$.

5.1.2 Image processing

The image reduction follows the principles outlined in Sects. 4.1 and 4.2. However, for technical reasons we deviated from this general version in several places – the differences are described below.

Pre-processing To remove the bias level in each frame we used the mean value of the overscan region from the corresponding chip. The flatfielding was carried out using a superflat, made from averaging 56 night time exposures; most of them blank fields (containing no bright stars leaving large extended haloes) and all well displaced from each other. The fringing that occurs in I -band exposures is also captured in this type of flat, and upon dividing the science exposures by the superflat the fringes were cleanly removed

¹IMCAT is developed by Nick Kaiser (kaiser@hawaii.edu), <http://www.ifa.hawaii.edu/~kaiser/imcat/>

(see below). To estimate the background level in the exposures, we used the heights of the minima of the sky level present to create a model for each individual frame. After subtraction the median sky level was set to zero.

As fringing is an additive effect and not a multiplicative one, ideally the fringes should be subtracted. Since we had no twilight flats available, standard defringing could not be performed. The photometric error introduced by division is negligible, as the amplitude of the fringes compared to the sky background after flatfielding was of the order of 2%. However, since fringing acts mostly on small angular scales, its treatment will affect the shapes of the small and faint background galaxies used for weak lensing. To investigate this we obtained a set of 10 archival images of the Deep3 field (Hildebrandt et al. 2006), taken with the Wide Field Imager at the 2.2m MPG/ESO telescope through their *I*-band filter. As the Deep3 field does not contain any massive clusters it is very well suited for this test. Two different coadded images were created. In the first case the data were flatfielded using twilight flats, after which a fringe model was created from the flatfielded data and subtracted. The second coadded image was processed in the same way as our CFHT data, i.e. the data were flatfielded and fringe-corrected by division of a night-time flat. We then measured the shapes (Sect. 2.3) of a common set of $\sim 12\,000$ galaxies in both images and created two mass reconstructions, using the same technique and smoothing scale as was done for Abell 1351 and Abell 1995 in Sect. 5.3. We found that the rms (root mean square) of the difference of the two mass maps is a factor of 2.5 smaller than the noise of the individual mass maps, mainly caused by the intrinsic ellipticities of background galaxies (eqs. (2.45) and (2.46)). The effect in our CFH12K data is much smaller, as the CFH12K *I*-band filter has a blue cut-on at around 730 nm and a red cut-off at 950 nm. The ESO *I*-band filter on the other hand opens at 800 nm and has no cut-off on the red side. Hence the fringing amplitude in the comparison data set from ESO is up to 5 times higher than that from CFHT. We thus conclude that our analysis of Abell 1351 and Abell 1995 is not affected by our fringe correction.

Masking The CFH12K mosaic contains some bad pixels and columns, in particular two of the CCDs suffer from this. By using Nick Kaiser’s ready made CFH12K masks² as global masks, all bad areas were ensured to be ignored. An additional patch of 219 bad columns in CCD00 was also added to the global masks. We did not make further individual masks for each exposure, as most spurious detections were filtered out during the astrometric calibration and by using a median during coaddition. Suspicious objects in the final catalogue were in addition rejected by visual examination.

Astrometric calibration Wide-field data typically do not have a simple relation between the sky coordinates and those of the detector. A mapping from pixel coordinates onto a planar projection of the sky needs therefore be performed. We solved for this through a series of steps, which differ from the standard procedure detailed in Sect. 4.2.3.

First, all objects in each exposure were detected and aperture photometry (Sect. 4.3.3) carried out. By plotting r_g (\sim half-light radius; the radius within which half the object’s luminosity is contained) vs. instrumental magnitude of each exposure’s objects (left plot of Fig. 5.1), we extracted the moderately bright, non-saturated stars suitable for deriving an

²<http://www.ifa.hawaii.edu/~kaiser/cfh12k/masks/>

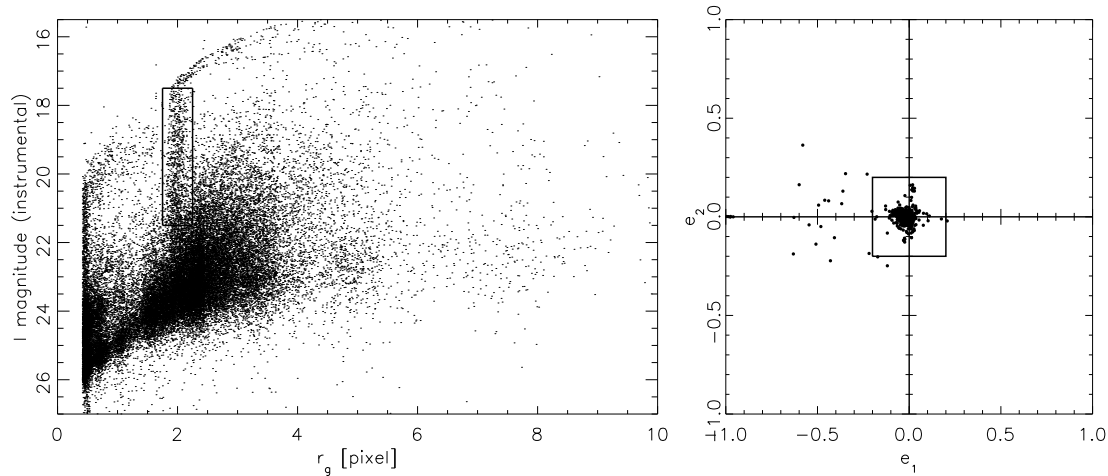


Figure 5.1: Example of size-magnitude (*left*) and weighted ellipticity parameter diagram (*right*) for an arbitrary exposure. The moderately bright, non-saturated stars were chosen from within the rectangles and utilised in the astrometric solution, the right plot containing only the stars chosen from the left plot. The stars utilised in the PSF corrections (Sect. 5.2) were selected the same way.

astrometric solution (i.e. those inside the rectangle). Their weighted ellipticity parameters, $e_{1/2}$ (defined in eq. (2.38)), were in addition plotted and eyeballed, and the main clustering of objects were selected to ensure a catalogue containing purely stars (right plot of Fig. 5.1).

Left with star catalogues we then computed the transformation parameters needed using information from the USNO-B1.0 catalogue (Monet et al. 2003). However, as many of the USNO-B1.0 stars were saturated in our images, we extended our reference catalogue by detecting more sources in FITS files derived from the Digitized Sky Survey³. We matched the target catalogues to the reference catalogue, solving for a set of low-order spatial polynomials mapping the images onto each other, by repeatedly refining the least squares minimisations using outlier rejection.

Final master image and object catalogue Sensitivity variations between chips (due to quantum efficiency variations) and differential extinction between exposures (f.ex. through clouds) were corrected. The extinction corrections between the exposures were very small, typically ~ 0.01 magnitudes, whereas the zero-point offsets between the chips were ~ 0.1 magnitudes. As an accurate absolute photometric calibration was not necessary for the present work, we adopted standard Landolt magnitude zeropoints (Sect. 4.3.3).

The coaddition was done after magnitude corrections were applied to the data. In addition cosmic rays were masked out, before the median image was computed and the background flattened. A master object catalogue was created, where each object's WCS (world coordinate system) coordinates were calculated from the astrometric solution, and their ellipticity parameters computed (Sect. 2.3.1). Finally we masked out false detections by overplotting the objects onto the image, hence ensuring a final object catalogue free from spurious detections.

³<http://archive.stsci.edu/dss/index.html>

5.2 Shear measurements

We utilised the KSB+ (Kaiser et al. 1995; Luppino & Kaiser 1997; Hoekstra et al. 1998) method to recover the image shear in our data. KSB+ inverts the effects of PSF smearing and anisotropy to recover the true shear of an object image (Sects. 2.2 and 2.3).

The ellipticities of the stars were fitted to a sixth-order Taylor series expansion in order to model the corresponding ellipticities at the galaxy positions. When comparing mass and noise maps (Sect. 5.3) for fits of different orders, there was little change with the order of fit. Over the whole field, 410 and 530 stars were used in the fitting process for Abell 1351 and Abell 1995 respectively. Figure 5.2 shows the ellipticities of the stars before and after PSF corrections.

Hoekstra et al. (1998) showed that estimating the PSF dilution for each individual galaxy introduces additional noise. We therefore followed their approach by determining p^γ (eq. (2.66)) as a function of magnitude and galaxy size. We determined the median p^γ within 15 bins in an r_g -magnitude diagram, where the central 4 bins contain ~ 4000 galaxies/bin and the outer ones ~ 200 galaxies/bin. We then computed one correction factor for each bin using eq. (2.66) and applied this to all galaxies within the corresponding bin.

The faintest and smallest galaxies are more affected by seeing than the larger galaxies, giving them a poorer shape determination and a high correction factor. Such galaxies are therefore of less importance. To account for this, a normalised weight,

$$w_i \propto \left(\frac{\sigma_{e_i}}{\langle p^\gamma \rangle_i} \right)^{-2}, \quad (5.1)$$

was calculated for the galaxies in bin i . Here, σ_{e_i} is the observed dispersion of galaxy ellipticities (similar to eq. (2.46)).

5.3 Mass reconstruction

Section 2.5 presents a general description on two-dimensional mass reconstruction. We selected background galaxies with $6 < S/N < 100$ for the creation of our mass maps. These maps were constructed from the galaxies' shapes using the finite-field inversion method presented by Seitz & Schneider (2001; SS01). This method calculates a smoothed shear field on a grid using a modified Gaussian filter. The algorithm then iteratively computes a quantity $K(\theta) \equiv \ln[1 - \kappa(\theta)]$ which is determined up to an additive constant due to the mass-sheet degeneracy (Sect. 2.2.3). We could break this degeneracy by assuming that the average convergence vanishes along the border of the wide field of view. The width of the Gaussian term in the filter was set to $2'0$, resulting in an effective smoothing length of about $1'6$.

In order to evaluate the noise of the mass maps, we computed 2000 mass maps for each cluster based on randomised galaxy orientations, keeping their positions fixed. As the cluster lens signal increases the ellipticities of galaxies, this would lead to an overestimation of the noise at the cluster position. We roughly corrected for this effect by subtracting the expected SIS tangential shear signal, determined from the clusters' known velocity dispersions (see Table 5.1). Since the singularity of the SIS can lead to overly large

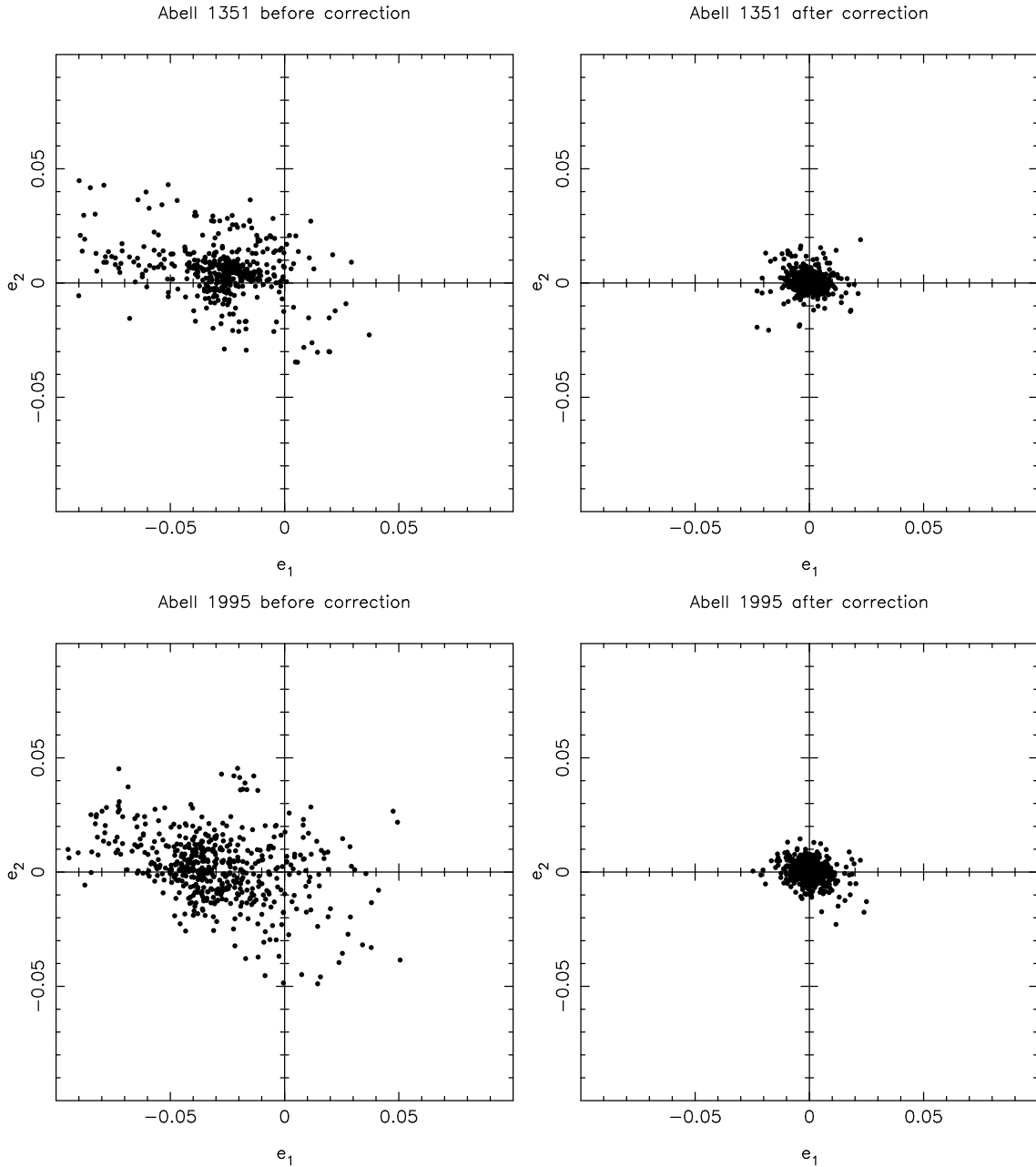


Figure 5.2: Ellipticities of the stars in the field of Abell 1351 (top) and Abell 1995 (bottom) before and after corrections for PSF anisotropies. The stars initially had systematic ellipticities up to $\sim 7 - 9\%$ in one direction. The PSF corrections reduced these effects to typically $< 1.5\%$.

corrections close to the cluster centre, we limited the maximum correction factor allowed to 0.5 in each ellipticity component. This affected less than 5 galaxies in both fields. The true mass maps were then divided by the noise maps obtained from the randomised mass maps to create the S/N-maps in Fig. 5.3.

Abell 1351 and Abell 1995 are detected with a S/N of 5.3 and 5.2, respectively. Upon integrating the κ maps within $r_{200} = 1.69 h_{70}^{-1}$ Mpc ($1.50 h_{70}^{-1}$ Mpc) for Abell 1351

5.3. Mass reconstruction

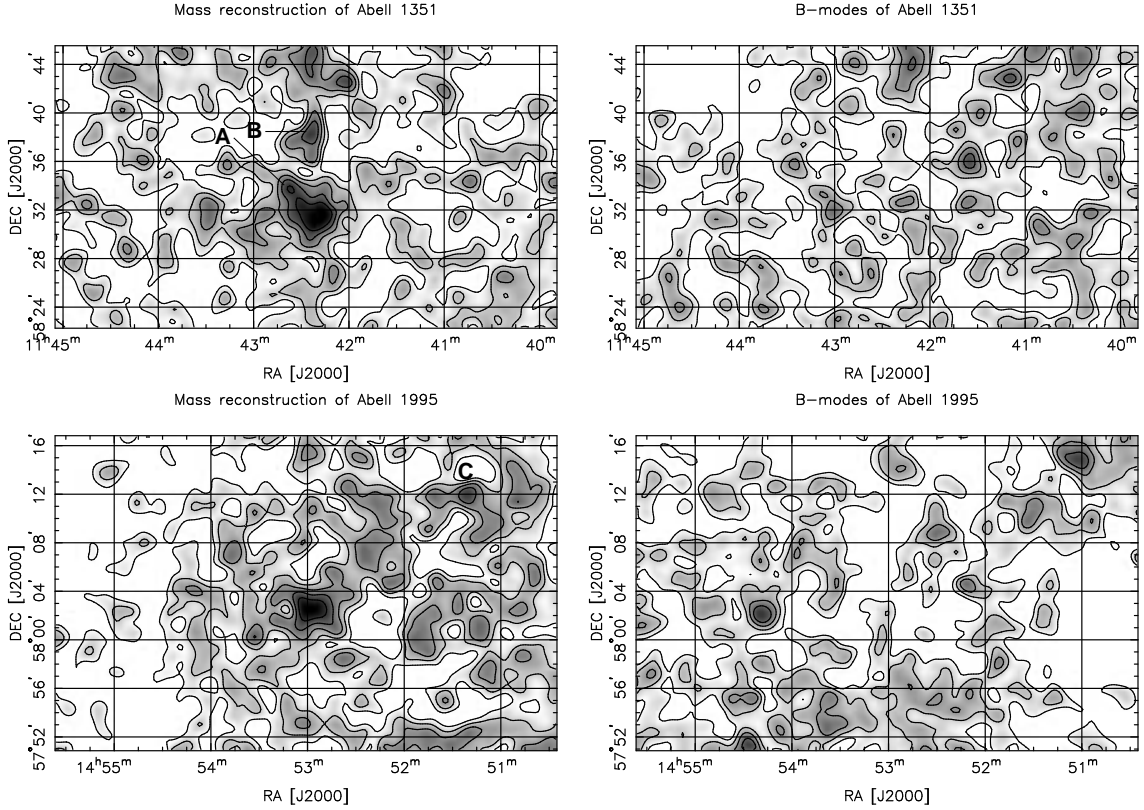


Figure 5.3: The projected surface mass densities and B modes for both clusters in the full CFH12K field, using a finite-field mass reconstruction. The maps show the S/N of the clusters, with contours starting at 0σ and increasing in steps of 1σ . The clusters are detected at significance levels of 5.3σ (Abell 1351) and 5.2σ (Abell 1995). The lower peaks A, B, and C have S/N-ratios of 4.2, 3.8, and 3.8, respectively. The effective smoothing length of the reconstructions is $1''.6$. The highest B-mode peak of Abell 1351 (Abell 1995) has a S/N-ratio of 3.5 (3.9).

(Abell 1995), we find total masses of $M_{2D}(< r_{200}) = (11.7 \pm 3.1) \times 10^{14} h_{70}^{-1} M_{\odot}$ and $M_{2D}(< r_{200}) = (10.5 \pm 2.7) \times 10^{14} h_{70}^{-1} M_{\odot}$ for the clusters, respectively. The r_{200} radii have been taken from what we consider to be the best NFW fits to the data (see Table 5.1 and Sect. 5.4.2), while the errors were determined from integrating the same areas in the 2000 noise maps.

The B modes in both cluster fields are shown in Fig. 5.3, computed by repeating the κ reconstruction with all galaxies rotated 45° (Crittenden et al. 2002). Such B-mode maps serve as a test for systematics in the lensing data. Provided that the data are free from systematics and that the noise (intrinsic ellipticities) is Gaussian, the B-mode maps should be consistent with Gaussian noise. Given the effective filter scale of $1''.6$, about 380 independent peaks can be placed in the CFH12K field. Thus one would expect 1.1 noise peaks above 3σ for a Gaussian random field. A more realistic estimate comes from the 2000 randomisations as these are based on the real ellipticity and spatial distribution. We expect 1.4 (1.6) such peaks for Abell 1351 (Abell 1995). In the real B-mode maps we find 3 peaks for each of the clusters. This is insignificant, as in our randomisations at least 3 such peaks appear per field in 20% of the cases. In case of Abell 1995 the highest B-mode

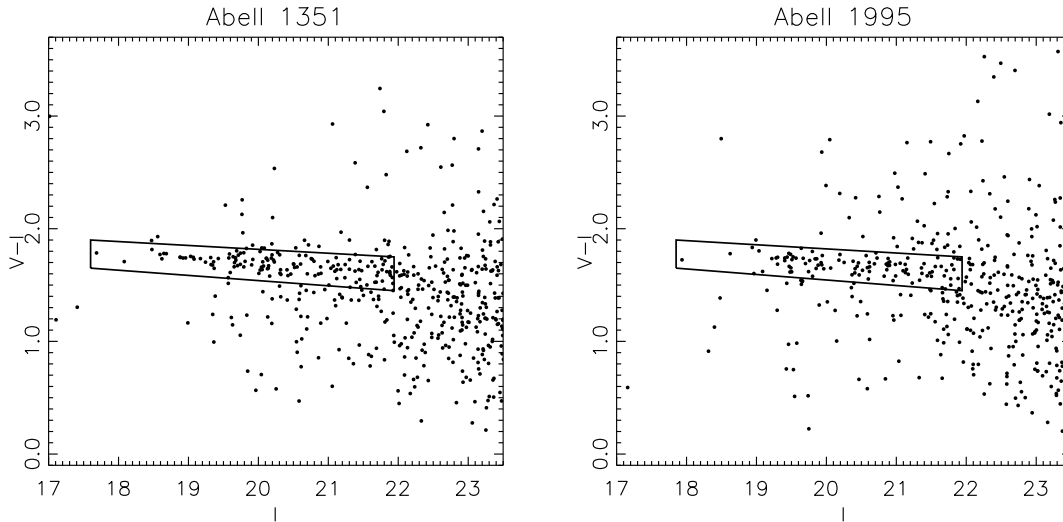


Figure 5.4: Colour-magnitude diagrams for Abell 1351 (left) and Abell 1995 (right), where only galaxies within $3'$ from the BCG are plotted. The red sequence can be seen for $1.4 < V - I < 1.9$ for both clusters, the box indicating our selection criteria.

peak has a significance of 3.9σ . Its B modes appear generally somewhat larger than for Abell 1351, which has no B-mode peaks higher than 3.5σ .

5.3.1 Mass and galaxy density distributions

In order to compare the surface density maps with the distribution of cluster galaxies, we extracted the red sequence (see Sect. 3.1 and also e.g. Gladders & Yee 2000) and investigated the distribution of the selected galaxies. To match the I -band and the V -band images, we resampled both data sets to the same pixel scale, resulting in a common area of $18'.5 \times 18'.5$ on the sky. The V -band image seeing is around $0''.7$, and thus consistently better than in the I -band. The V -band data were therefore convolved to match the seeing in the I -band data, $0''.95$ for Abell 1351 and $1''.15$ for Abell 1995. Aperture photometry was carried out using `SExtractor` (Bertin & Arnouts 1996) in double-image mode. The deep I -band images served as the detection image, providing us with a target list with defined coordinates. At these positions we integrated the flux in a $3''$ wide aperture in each of the V - and I -band images. Plotting the galaxies in a colour-magnitude diagram will then in principle provide enough information to separate the red early-type cluster members from the other galaxies.

Each cluster's red sequence did not clearly stand out from the $V - I$ vs. I diagram when all objects were plotted. However, selecting only galaxies within a radius of $3'$ of the brightest cluster galaxy (BCG) for the colour-magnitude diagram (Fig. 5.4) enabled us to detect a red sequence at $1.4 < V - I < 1.9$ for both clusters. The selection criteria indicated by the box in each plot were then applied to the entire object catalogue. The number density of the galaxies selected was calculated as a function of position and overplotted onto the central $9'$ of the mass maps (Fig. 5.5).

The number density maps were normalised by the fluctuation measured in the field

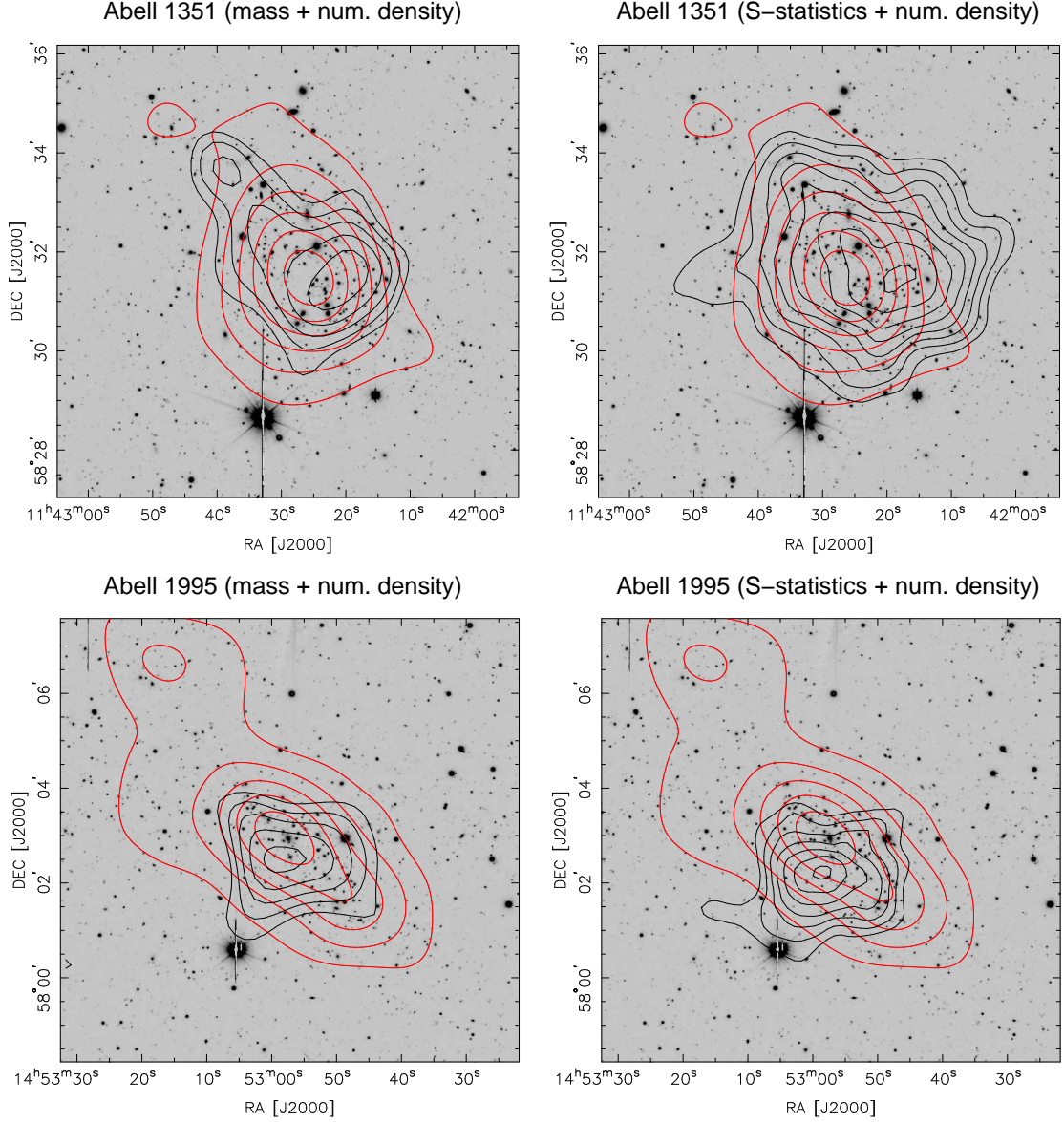


Figure 5.5: The black contours show the mass reconstruction (left) and S -statistics (right) for Abell 1351 (top) and Abell 1995 (bottom). The contours start at the 3.0σ level and increase in steps of 0.5σ . Overplotted in red contours are the number densities of galaxies selected with the red sequence method, normalised by the rms fluctuation in the field outside the clusters. These contours are isodensity contours, starting at 3σ going in steps of 3σ , and peak at 15σ and 10σ , respectively.

outside the clusters. The centres of Abell 1351 and Abell 1995 are then detected with 15.5σ and 9.9σ significance, respectively. The positional offsets between mass centres, BCGs, and centres of galaxy density distributions are in the range of $30'' - 55''$ for both clusters, and are due to noise in the mass maps. Changing the width of the Gaussian kernel in the finite-field reconstruction algorithm shows that the peak centres can drift by up to $20''$ from the mean position. These offsets are consistent with other results in the literature,

such as Clowe et al. (2006), who observed offsets of the order of $10''$ between the lensing peaks (of higher S/N than ours) and the optical centres of the bullet cluster. Positional offsets of $50''$ are common in the sample of 70 shear-selected clusters by Schirmer et al. (2007).

The cluster galaxy distribution resembles well the mass distribution in the central part of Abell 1351. It extends significantly towards the northeast, a feature also seen in the mass map where we find a local maximum which we refer to as peak A (see Sect. 5.3.2). The galaxy distribution of Abell 1995 appears elliptical and elongated in the northeast-southwest direction. This property is not reflected in the mass map where the peak is of rather circular appearance.

To check the integrity of our mass reconstructions further, we applied the peak finder (S -statistics) developed by Schirmer et al. (2007), which is a generalised version of the aperture-mass statistics introduced by Schneider (1996). This method detects areas of enhanced tangential shear in a grid projected onto the sky, normalised by a local noise estimate obtained from the data. Hence the S -statistics directly returns the S/N of the lensing detection. The filter function Q is chosen such that it approximates the expected NFW shear profile, maximising the S/N. At a grid point θ_0 the S -statistics is evaluated as

$$S(\theta_0) = \frac{\sqrt{2} \sum_i \varepsilon_{ti} w_i Q_i}{\sqrt{\sum_i |\varepsilon_i|^2 w_i^2 Q_i^2}}, \quad (5.2)$$

where Q_i is the value of the filter function at the position of a galaxy i while centred on θ_0 , w_i represents the weight of each individual galaxy, and ε_i (ε_{ti}) is the galaxy's (tangential) ellipticity (Sect. 2.3.1).

Since the S -statistics uses a filter function that mimics the tangential shear profile of galaxy clusters it is well suited for detecting mass concentrations. With this method we recover Abell 1351 at the 7.0σ level in a $10'$ wide filter, and Abell 1995 with 6.1σ for a $7'$ filter. The filter shape parameter (Schirmer et al. 2007) was chosen as $x_c = 0.2$ in both cases. We find that the S -statistics is elongated in the same way as the mass reconstruction for Abell 1351, extending towards peak A. We evaluate the significance of this possible substructure in the following.

5.3.2 Lower mass peaks in the fields

In the mass reconstructions two neighbouring peaks A and B are detected around Abell 1351, and another one (peak C) in the field of Abell 1995. Their S/N-ratios are 4.2, 3.8, and 3.8, respectively. We used the 2000 noise randomisations for each field to find that the probability of a noise peak higher than 4.2σ (3.8σ) in the field of Abell 1351 is 0.8% (6.8%), respectively. The corresponding probability for peak C in the Abell 1995 data is 7.0%. These are somewhat higher than what would be expected from purely Gaussian noise.

Hence the only significant substructure detected in the mass reconstructions is peak A near Abell 1351. Looking at the contours in the upper right panel of Fig. 5.5, one can see that the S -statistics trace this structure as well at the $4.5 - 5.0\sigma$ -level. We note that we recovered this substructure over a broad range of filter scales in the S -statistics and hence think that it is a real feature in the mass distribution of Abell 1351.

Out of the broad range ($1' - 15'$) of filters probed with the S -statistics, peak B is detected only once with 4.0σ in the $4'$ wide filter for $x_c = 0.5$. It has the typical characteristics of the dark peaks found by Schirmer et al. (2007), i.e. it is not associated with any overdensity of galaxies. It is therefore most likely a noise peak.

In the Abell 1995 field we could not detect any other significant peaks using the S -statistics. Since the B modes for those data show a maximum of 3.9σ near peak C (at 3.8σ), we consider it to be a noise peak. As it also lies outside the area covered by the V -band, we could not check for overdensities of red galaxies at this position.

5.3.3 Giant radio halo in Abell 1351

Using archival Very Large Array data at 1.4 GHz, Giacintucci et al. (2009) detect a giant radio halo centred on the main mass peak of Abell 1351. Its elongation and overall size agrees well with the brightest X-ray region, detected in archival ROSAT High Resolution Imager (HRI) data. The radio emission peak also coincides with the X-ray peak.

A faint filament of radio emission stretches towards the northeast, extending in the direction of peak A in our mass reconstruction, further confirming the existence of a secondary mass substructure in this cluster. Based on the work of Cassano & Brunetti (2005), Giacintucci et al. (2009) also find that the radio emission observed is expected to be found in merger systems with a total virial mass of $M_{\text{vir}} \simeq 1.2 \times 10^{15} M_{\odot}$, consistent with our results of $M_{2D}(<r_{200}) = (1.17 \pm 3.1) \times 10^{15} h_{70}^{-1} M_{\odot}$.

5.4 Modelling the lensing data

The mass of a galaxy cluster can be estimated by comparing observed distortions in the background galaxies to those predicted by theoretical density profiles (Sect. 2.4). Using χ^2 -minimisations of SIS and NFW models we first determined the best fit parameters and then calculated the cluster masses.

The theoretical profiles are both spherically symmetric. We therefore averaged the tangential reduced shear, $g_t = \gamma_t / (1 - \kappa)$ (for $r > \theta_E$, where θ_E is the Einstein radius, eq. (2.73)), in 17 radial bins around the cluster centre. The bins are logarithmically spaced, covering the entire field of view, and starting at $r_{\text{min}} = 150''$ to avoid the large contamination from cluster galaxies close to the centre of the field (see also Sect. 5.4.1). To determine the cluster centre, we tested three different positions. First we adopted the peak location in the mass reconstructions generated (Sect. 5.3). These coincide with the centres of the S -statistics. Second the BCG served as cluster centre, and third we tried the centre of the galaxy density of each cluster. As the latter coincides with the BCG for Abell 1995, only two positions were tested for this cluster. We also considered strong lensing features, but found that they do not offer further insight in this respect (see Sect. 5.5.3a for details). We calculated $\langle g_t \rangle_i$ for each radial bin i and compared them to the theoretical values at the average radius of each bin, $\langle r \rangle_i$ (eq. (2.48)).

When calculating the mass of a cluster, the relative distance of the background galaxies and the lensing cluster is required. As we have no specific information about the redshifts of the background galaxies, the distances had to be estimated statistically. By using the photometric redshift (Sect. 4.4.1) distribution of corresponding faint galaxies from the Hubble Deep Field North (Fernández-Soto et al. 1999), we estimated the average

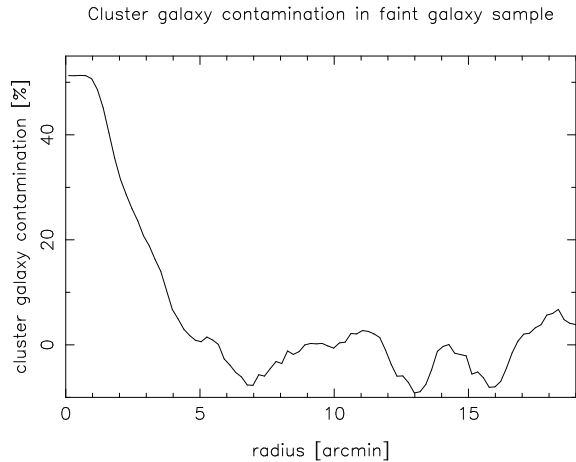


Figure 5.6: Percentage of cluster galaxies in the faint galaxy catalogue of Abell 1351 (that of Abell 1995 is very similar). Because the projected density of cluster galaxies is assumed to equal zero at the edge of the field, the cluster galaxy contamination was set to zero here by subtracting the median value outside the central area of the image (the large field of view makes this a well-working approximation). Our fitting procedure starts at $r_{\min} = 150''$ to avoid the large cluster galaxy contamination at the centre.

$\beta \equiv D_{\text{ds}}/D_s$, where D_{ds} is the angular diameter distance between the lens and the source and D_s between the observer and the source (see also Sect. 1.7). The empirical relation

$$\langle \beta \rangle = -1.21z_d + 0.91 \quad (5.3)$$

is derived for a Λ CDM cosmology analogously to the equation of Dahle et al. (2002) for an Einstein-de Sitter universe (where $\Omega_m = 1$, $\Omega_\Lambda = \Omega_r = 0$), which also accounts for the redshift distribution of the source galaxies. Here z_d denotes the redshift of the lens.

5.4.1 Cluster contamination and magnification depletion

At small projected radii from the cluster centre our faint galaxy sample will contain cluster galaxies in addition to background galaxies. We could not discriminate faint cluster members from lensed field galaxies using $V - I$ colours, hence the sample of presumed lensed background galaxies remained contaminated, leading to a systematic bias of the shear measurements towards lower values. In order to quantify this contamination, we determined the overdensity of galaxies in the background catalogue at the cluster position compared to the mean density in the field (an example is shown in Fig. 5.6 for Abell 1351). A contamination rate of 50% was found for both cluster centres, vanishing for radii larger than about $4' - 5'$.

The cluster galaxy contamination is in fact even higher than stated above, as magnification depletion leads to a reduced number density of lensed galaxies in the I -band near the cluster centre. We found, however, that this effect can be neglected in our case. From the smoothed convergence (Sect. 5.3) and reduced shear fields we estimated the magnification using eq. (2.32). We found very similar magnifications for both clusters, being 1.65 at the centre and becoming indistinguishable from the noise ($\sigma_\mu \sim 0.15$) for radii larger than $\sim 3'$. The depletion of the number density of lensed galaxies is $\propto \mu^{2.5s-1}$, with $s = 0.15$ in red filters (see e.g. Narayan & Bartelmann 1996). At the cluster centres the number densities are thus reduced by a factor of ~ 0.73 , and at a radius of $1.5'$ magnification depletion becomes indistinguishable from the natural fluctuations in the distribution of field galaxies. Magnification depletion hence only affects the innermost ~ 0.3 Mpc ($100''$) of the clusters and can be neglected since we compare the tangential shear profiles to models only for radii larger than 0.5 Mpc (see Fig. 5.7).

In order to correct for the contamination by cluster galaxies, we modified the theoretical shear values. The reason for adjusting the theoretical values rather than the measured values is that this method is considerably easier to implement. The correction factors were determined in radial bins of logarithmic spacing. One correction value was then calculated for each of the 17 bins in which $\langle g_t \rangle$ had been measured. By assuming the edges of the field to be approximately free from cluster galaxies, the outermost correction factor could be set to 1 to mimic contamination-free boundaries of the field. Finally the best fit was found using χ^2 -minimisations.

5.4.2 Fitting the SIS and the NFW profiles

Once the cluster centre is determined, the only free parameter of the SIS profile is the velocity dispersion, σ_v (eq. (2.76)). The best fit of the SIS profile is determined by χ^2 -minimisation for a range of σ_v values, the results being shown in Table 5.1. The mass estimate, M_{SIS} , for this profile is calculated at r_{200} (the radius inside which the mean mass density of the cluster equals 200 times the critical density of the Universe) found in the NFW fitting with two free parameters utilising the same cluster centre. The SIS mass model is described in detail in Sect. 2.4.2.

An introduction to the gravitational lensing properties of the NFW mass density profile is given in Sect. 2.4.3. We derived the best fit parameters for different values of the concentration parameter, c , ranging from 0.1 to 24.9 in steps of 0.1. With the cluster centre fixed, the NFW profile has two free parameters, r_{200} and c . We fitted our shear measurements to this profile twice; first by keeping c fixed and varying only r_{200} to find our best fit, and second by varying both parameters. The best fit parameters were determined by minimising χ^2 in both cases. Based on N-body simulations of dark matter halos (Sect. 3.4.3), Bullock et al. (2001) derived relations for the mean value of c as a function of redshift and mass for different cosmologies. For a halo of mass $M_{\text{vir}} = 8 \times 10^{14} M_{\odot}$, the relation yields

$$c = \frac{5.8}{1.194(1 + z_d)} \quad (5.4)$$

(where $r_{200} = 1.194 r_{\text{vir}}$ for a Λ CDM cosmological model). As this mass is close to the mass estimates of Abell 1351 and Abell 1995 (Sect. 5.3), the weak mass dependence of c can be disregarded. In the second fitting process both r_{200} and c were altered, creating a grid of c, r_{200} -values. The best fit r_{200} was first determined for each value of c , then the best fit c was found. The results are given in Table 5.1. We could not find an upper limit for the concentration parameter of Abell 1351, independent of the cluster centre chosen. The same holds for Abell 1995 in case the BCG is chosen as the centre. We discuss this in more detail in Sect. 5.5.2.

As an example we display the reduced tangential shear as a function of radius using the SS01 κ maps as cluster centre, see the left diagrams of Fig. 5.7. The measured values of $\langle g_t \rangle$ are given together with the best fit models of the SIS and NFW profiles. Judging from the diagrams alone, the NFW profile letting both c and r_{200} vary seems to represent the best fit to the clusters. However, the χ^2/Dof values given for each model in Table 5.1 show that the differences between the models are not statistically significant. The differences emerging from different cluster centres seem to have a higher impact. The right diagrams

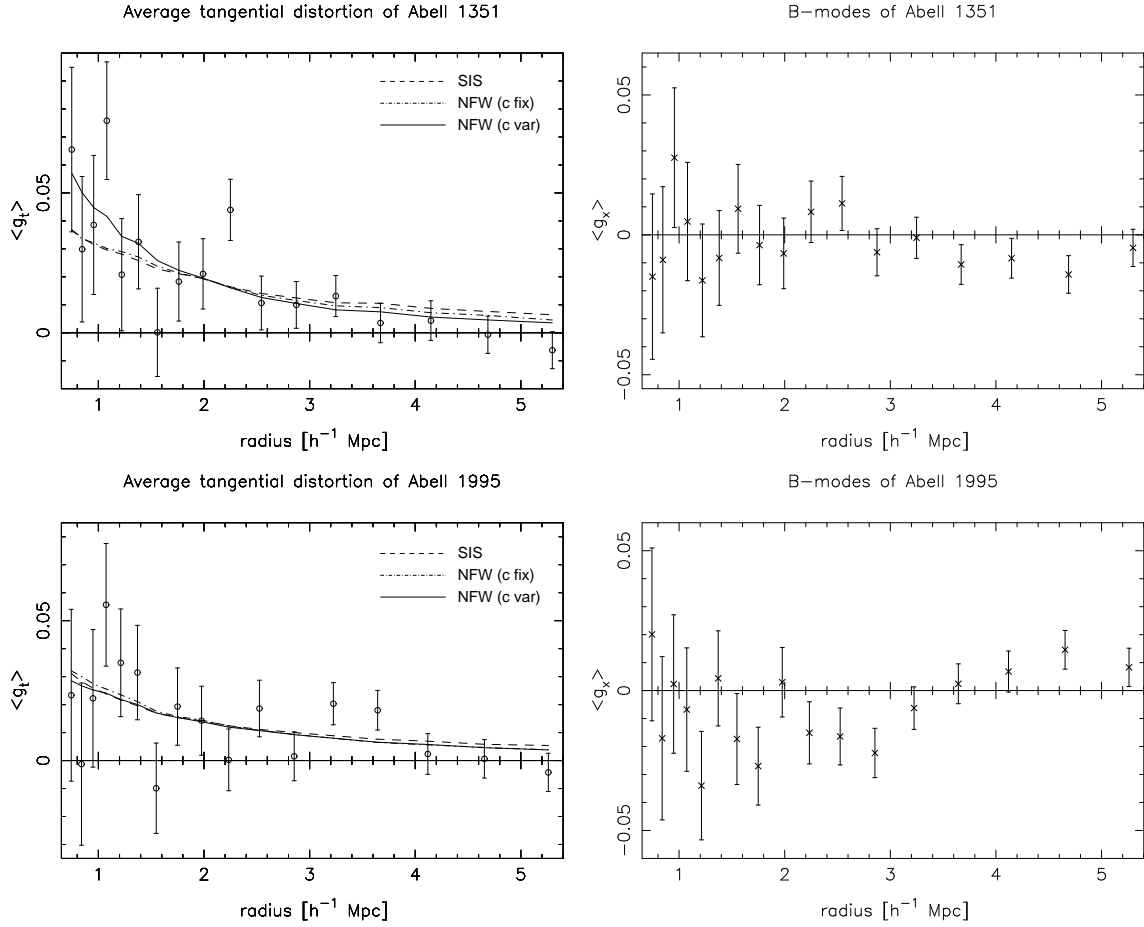


Figure 5.7: *Left*: Reduced tangential shear as a function of radius for Abell 1351 (top) and Abell 1995 (bottom) using the SS01 κ map as cluster centre (the other centres show very similar figures). The averaged gravitational lensing distortions of background galaxies are shown as points with 1σ error bars. The lines indicate the best fit models; the dashed line representing the SIS profile, the dot-dashed line the NFW profile keeping c fixed, and the solid line the NFW profile with two free parameters. It should be noted that these lines represent the theoretical values *after* modifications from cluster galaxy contamination are applied to each bin independently of each other (see Sect. 5.4). *Right*: Cross-component of the reduced tangential shear as a function of radius for Abell 1351 (top) and Abell 1995 (bottom). This signal should disappear if $\langle g_t \rangle$ is caused by lensing only, and it is seen that the measurements are close to zero for both clusters.

of Fig. 5.7 show the B modes of both clusters, i.e. the cross-component of $\langle g_t \rangle$ as a function of radius. Both measurements are consistent with zero.

5.5 Discussion

X-ray studies show Abell 1351 to be a galaxy cluster exhibiting significant dynamical activity and undergoing a major merger event (Allen et al. 2003), which indicates a cluster still in its formation phase. Analyses assuming a relaxed cluster (Sect. 3.3) will hence naturally differ from weak lensing analyses, where no assumption is made about the dy-

Table 5.1: Results from fitting theoretical density profiles to the measured shear values, where each subheadline indicates the cluster centre around which the fitting was done.

BCG	Abell 1351	Abell 1995	SS01^a κ map	Abell 1351	Abell 1995
SIS [χ^2 / Dof]	1.60	1.70	SIS [χ^2 / Dof]	1.44	1.14
θ_E ["]	16.2 ^{+2.3} _{-2.4}	12.2 \pm 2.6	θ_E ["]	15.0 ^{+2.5} _{-2.3}	12.0 ^{+2.5} _{-2.8}
σ_v [km s ⁻¹]	1040 ⁺⁷⁰ ₋₈₀	900 ⁺⁹⁰ ₋₁₀₀	σ_v [km s ⁻¹]	1000 \pm 80	890 ⁺⁹⁰ ₋₁₁₀
M_{SIS} [$10^{14}h_{70}^{-1}M_{\odot}$]	8.7 ^{+1.3} _{-1.4}	5.8 ^{+1.2} _{-1.3}	M_{SIS} [$10^{14}h_{70}^{-1}M_{\odot}$]	7.8 ^{+1.3} _{-1.4}	5.5 ^{+1.3} _{-1.4}
NFW [χ^2 / Dof]	1.32	1.64	NFW [χ^2 / Dof]	1.12	1.11
c	9.1 ⁺ _{-4.6}	5.2 ⁺ _{-3.4}	c	16.1 ⁺ _{-8.4}	3.0 ^{+9.0} _{-1.9}
r_{200} [h_{70}^{-1} Mpc]	1.73 ^{+0.13} _{-0.10}	1.53 ^{+0.17} _{-0.13}	r_{200} [h_{70}^{-1} Mpc]	1.69 ^{+0.10} _{-0.11}	1.50 ^{+0.20} _{-0.11}
M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	8.1 ^{+1.8} _{-1.4}	5.6 ^{+1.9} _{-1.4}	M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	7.5 ^{+1.3} _{-1.5}	5.3 ^{+2.1} _{-1.2}
NFW (fixed c)			NFW (fixed c)		
[χ^2 / Dof]	1.42	1.66	[χ^2 / Dof]	1.31	1.12
c	3.7	3.7	c	3.7	3.7
r_{200} [h_{70}^{-1} Mpc]	1.81 ^{+0.057} _{-0.16}	1.56 ^{+0.13} _{-0.16}	r_{200} [h_{70}^{-1} Mpc]	1.67 ^{+0.16} _{-0.086}	1.51 ^{+0.16} _{-0.14}
M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	9.4 ^{+0.9} _{-2.4}	5.9 ^{+1.5} _{-1.8}	M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	7.4 ^{+2.1} _{-1.1}	5.4 ^{+1.7} _{-1.5}
KS93^b κ map			Galaxy density		
SIS [χ^2 / Dof]	1.30	2.32	SIS [χ^2 / Dof]	1.17	
θ_E ["]	16.5 ^{+2.6} _{-2.4}	10.4 ^{+2.7} _{-2.6}	θ_E ["]	16.5 ^{+2.3} _{-2.4}	
σ_v [km s ⁻¹]	1050 \pm 80	830 ⁺¹⁰⁰ ₋₁₁₀	σ_v [km s ⁻¹]	1050 ⁺⁷⁰ ₋₈₀	
M_{SIS} [$10^{14}h_{70}^{-1}M_{\odot}$]	8.9 \pm 1.5	4.9 ^{+1.3} _{-1.4}	M_{SIS} [$10^{14}h_{70}^{-1}M_{\odot}$]	9.0 ^{+1.3} _{-1.5}	
NFW [χ^2 / Dof]	0.96	2.22	NFW [χ^2 / Dof]	0.88	
c	11.2 ⁺ _{-4.9}	0.9 ^{+2.1} _{-0.4}	c	11.3 ⁺ _{-6.7}	
r_{200} [h_{70}^{-1} Mpc]	1.74 ^{+0.11} _{-0.10}	1.50 ^{+0.17} _{-0.24}	r_{200} [h_{70}^{-1} Mpc]	1.76 ^{+0.10} _{-0.11}	
M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	8.3 ^{+1.6} _{-1.4}	5.3 ^{+1.8} _{-2.6}	M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	8.5 ^{+1.5} _{-1.7}	
NFW (fixed c)			NFW (fixed c)		
[χ^2 / Dof]	1.13	2.33	[χ^2 / Dof]	1.00	
c	3.7	3.7	c	3.7	
r_{200} [h_{70}^{-1} Mpc]	1.81 ^{+0.10} _{-0.13}	1.51 ^{+0.11} _{-0.23}	r_{200} [h_{70}^{-1} Mpc]	1.79 ^{+0.10} _{-0.11}	
M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	9.4 ^{+1.6} _{-2.0}	5.4 ^{+1.2} _{-2.5}	M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	9.0 ^{+1.5} _{-1.7}	
Integration of SS01 κ maps					
r_{200} [h_{70}^{-1} Mpc]	1.69	1.50			
$M_{2D}(< r_{200})$ [$10^{14}h_{70}^{-1}M_{\odot}$]	11.7 \pm 3.1	10.5 \pm 2.7			

^aSeitz & Schneider (2001).^bKaiser & Squires (1993).

Note that the KSB+ PSF correction tends to underestimate the shear by 10 – 15%, which in turn reduces the cluster masses up to 20%.

namical state of the cluster. One example is the virial analysis (Sect. 3.3.1) by Irgens et al. (2002), where an unusually high velocity dispersion of $\sigma_v = 1680_{-229}^{+340}$ km s⁻¹ was obtained for Abell 1351, based on radial velocity measurements of 17 cluster galaxies. Such a high velocity dispersion is not uncommon in merging systems. If for example two smaller clusters with low velocity dispersions fall towards each other along the line of sight with a velocity comparable to or higher than their σ_v , then a very large total σ_v would be inferred, with a correspondingly overestimated virial mass. The cluster CL0056.03 – 37.55 is a good example for such a system (see Schirmer et al. 2003).

Abell 1995 is, unlike Abell 1351, classified as a relaxed cluster in dynamical equilibrium (Pedersen & Dahle 2007). X-ray studies (Sect. 3.3.2) and virial analyses of this cluster are hence also more compatible with lensing studies (Patel et al. 2000; Irgens et al. 2002). The projected two-dimensional distribution of cluster galaxies in Abell 1995 is clearly elliptical (see Fig. 5.5), whereas the central lensing mass distribution is circular.

5.5.1 The mass estimates

The mass distributions of Abell 1351 and Abell 1995 were estimated assuming that the clusters follow spherically symmetric SIS or NFW profiles. Although an elliptical mass profile might yield more accurate cluster mass estimations, Dietrich et al. (2005) showed that the results from fitting a singular isothermal ellipse model depend strongly on the initial values chosen for the minimisation routines. We therefore decided not to fit this profile to our clusters.

Heymans et al. (2006) demonstrated in the shear testing programme that the KSB+PSF correction tends to systematically underestimate the shear values $\sim 10 - 15\%$. To measure how much this affects our data, we calculated an upper limit for our mass estimates by increasing the ellipticities with 15% and repeating the fitting process. We found that the underestimation of shear leads to an underestimation of the total cluster mass with a maximum of 20%, which is within the initial error bars. The concentration parameters did not change significantly by this boosting of ellipticities. Since we do not know by exactly how much our shear values are underestimated this was merely an attempt to quantify this effect on our data, and is not taken into account in the results presented.

Dahle et al. (2002) obtained weak lensing estimates of the cluster velocity dispersion of several clusters using an SIS model and assuming an Einstein-de Sitter universe. Their results were $\sigma_v = 1410_{-90}^{+80}$ km s⁻¹ for Abell 1351 and $\sigma_v = (1240 \pm 80)$ km s⁻¹ for Abell 1995, and do not agree with our results. However, there are several important differences in methodology between Dahle et al. (2002) and our work. As mentioned above, the assumed cosmological model is different. Also, the first paper approximated $g_t = \gamma_t$, whereas we use $g_t = \gamma_t/(1 - \kappa)$ in our fits and mass reconstructions. Finally, the shear estimator of Kaiser (2000) used by Dahle et al. (2002) is shown by Heymans et al. (2006) to have a non-linear response to shear. A re-analysis of the Dahle et al. (2002) data, taking all these effects into account, yielded new values of $\sigma_v = (1410 \pm 90)$ km s⁻¹ and $\sigma_v = (1000 \pm 100)$ km s⁻¹ for Abell 1351 and Abell 1995, respectively. Hence there still remains a systematic discrepancy between the results for Abell 1351, while the measurements for Abell 1995 agree within error bars.

A remaining difference between our work and Dahle et al. (2002) is the maximum radius, r_{\max} , to which the shear is measured, given by the field of view of the detector.

Changing r_{\max} in our Abell 1351 data to $550''$ (as this is the value used by Dahle et al. (2002)) led to $\sigma_v = (1240 \pm 105) \text{ km s}^{-1}$, consistent with the re-analysed Dahle et al. (2002) values within error bars.

Allen et al. (2003) used the Dahle et al. (2002) observations to obtain a weak lensing mass estimate applying the NFW model to a Λ CDM cosmology. Their results gave $M_{200} = 30.2^{+5.6}_{-4.9} \times 10^{14} h_{70}^{-1} M_{\odot}$ for Abell 1351 and $M_{200} = 14.4^{+3.3}_{-3.0} \times 10^{14} h_{70}^{-1} M_{\odot}$ for Abell 1995. These values are high compared to the results of this study. Allen et al. (2003) used a fixed concentration parameter in the fitting process, $c = 5$. Applying this value to our data yielded minimal changes in M_{200} . The discrepancies hence originate from Allen et al. (2003) utilising $r_{200} = 2.69^{+0.14}_{-0.19} h_{70}^{-1} \text{Mpc}$ and $r_{200} = 2.07^{+0.19}_{-0.14} h_{70}^{-1} \text{Mpc}$ (priv. comm.) for Abell 1351 and Abell 1995, respectively, as these values are larger than our best fit r_{200} values.

5.5.2 The concentration parameter

The mass density of a cluster with a low concentration parameter decreases slower when going to larger radii than for a cluster with a high c value (Wright & Brainerd 2000). Although unconstrained upwards, we find a lower limit of $c \geq 4.5$ for Abell 1351. As is also seen from the radial dependence of the shear in Fig. 5.7 (top left), the mass distribution of Abell 1351 concentrates around the cluster centre, indicating that its concentration parameter is significantly higher than that of Abell 1995. The values found for c of Abell 1995 (see Table 5.1) suggest that its mass is spread more evenly to larger radii, which is also seen in Fig. 5.7 (bottom left).

From their aperture mass calculations Dahle et al. (2002) found that most of the mass of Abell 1995 is contained within $r \sim 0.9 h_{70}^{-1} \text{Mpc}$ ($\sim 200''$). The mass of Abell 1351 shows the opposite behaviour, increasing evenly with radius, even at large radii. These results are contrary to our conclusions. As measurements at large radii are certain to include additional information not recognised close to the cluster centre, these discrepancies are likely explained by the difference in field-size between the two studies. By mapping the mass distribution towards a radius more than twice the size as that of Dahle et al. (2002), our results are better constrained. Further bias also arises from the measurements of Dahle et al. (2002) starting from an inner radius of $r_{\min} = 50''$ ($\sim 0.37 h_{70}^{-1} \text{Mpc}$), where we consider the cluster galaxy contamination to be very high, in addition to not including any correction for this contamination.

Our shear values are measured from a radial cut-off, $r_{\min} = 150''$, to avoid the large cluster galaxy contamination present at small radii. Because c is estimated from the scale radius, $r_s = r_{200}/c$, it is desirable to include r_s in the measurements ($r_{\min} < r_s$) in order to obtain an accurate estimate of the concentration parameter. If this is not the case, c is basically unconstrained.

This appears to be the case for Abell 1351, explaining why we were not able to derive an upper limit for its concentration parameter. Letting $r_{\min} = 150''$, we ensured a cluster galaxy contamination $< 25\%$ at this inner radius. However, as the c parameter appears unconstrained under this condition, we reduced r_{\min} in an attempt to obtain clearer results. The problem then arising was the increasing contamination of cluster galaxies. Looking at Fig. 5.6 we see that at $r = 120''$ the cluster contamination is $\sim 32\%$, and at $r = 100''$ it equals $\sim 40\%$. Though this contamination is accounted for during the fitting process, the

Table 5.2: Results from varying the inner radius from where the shear values of Abell 1351 are measured.

r_{\min} [$''$]	r_s [$''$]	c	r_{200} [$h_{70}^{-1}\text{Mpc}$]	M_{200} [$10^{14}h_{70}^{-1}M_{\odot}$]	No. of galaxies	χ^2/Dof
100	76^{+56}_{-35}	$4.9^{+3.6}_{-2.2}$	$1.76^{+0.057}_{-0.13}$	$8.5^{+0.8}_{-1.9}$	15 630	1.11
110	61^{+48}_{-29}	$6.0^{+4.7}_{-2.8}$	1.71 ± 0.10	7.9 ± 1.4	15 582	1.45
120	54^{+60}_{-26}	$6.8^{+7.6}_{-3.3}$	$1.71^{+0.10}_{-0.086}$	$7.9^{+1.4}_{-1.2}$	15 529	1.31
130	60^{+83}_{-26}	$6.2^{+8.5}_{-2.6}$	$1.76^{+0.057}_{-0.11}$	$8.5^{+0.8}_{-1.7}$	15 482	1.33
140	$37^{+\infty}_{-20}$	$10.0^{+\infty}_{-5.5}$	$1.73^{+0.11}_{-0.13}$	$8.1^{+1.6}_{-1.8}$	15 428	1.16
150	$33^{+\infty}_{-15}$	$11.2^{+\infty}_{-4.9}$	$1.74^{+0.11}_{-0.10}$	$8.3^{+1.6}_{-1.4}$	15 358	0.96

contamination correction is still vulnerable to fluctuations in the projected galaxy density caused by foreground and/or background structures.

Table 5.2 presents the results from letting $100'' \leq r_{\min} \leq 150''$ for Abell 1351 (with the Kaiser & Squires 1993 κ map peak as cluster centre, see Sect. 5.5.3). It is seen that whilst c is decreasing with smaller r_{\min} , r_{200} and M_{200} remain stable for different r_{\min} . Also worth noticing is that for $r_{\min} \leq 130''$, c becomes constrained. However, as $r_{\min} > r_s$ for the different starting radii, we cannot obtain further conclusions from these results. As r_s is even smaller for Abell 1995, we did not repeat this test for the cluster. Dietrich et al. (2005) experienced similar problems when attempting to determine the concentration parameter for Abell 222 and Abell 223, concluding that obtaining a reliable c from weak lensing data only is difficult, if not impossible.

5.5.2a Best fit concentration parameter

Bullock et al. (2001) presented dark matter halo simulations, attempting to find a “best fit concentration parameter” applicable to all types of halos. They found that for halos of the same mass, the concentration, $c_{\text{vir}} \equiv r_{\text{vir}}/r_s$, can be given by $c_{\text{vir}} \propto (1 + z_d)^{-1}$. This is contrary to earlier beliefs that c_{vir} does not vary much with redshift. Numerically simulated massive clusters typically have concentration parameters $\sim 4 - 5$ (Bullock et al. 2001). This is within the limiting values for both clusters, although looking at Fig. 5.7, the outcome from varying c seems to better follow the shear values of Abell 1351.

There exists several examples of high concentration parameters in the literature. Kneib et al. (2003) found $c = 22^{+9}_{-5}$ for the central mass concentration of the cluster Cl 0024 + 1654. Gavazzi (2005) concluded on $c = 11.73 \pm 0.55$ for MS2137 – 23, while Broadhurst et al. (2005) found $c = 13.7^{+1.4}_{-1.1}$ for Abell 1689. Limousin et al. (2007) presented a thorough discussion of the different concentration parameters derived for Abell 1689 in the literature, concluding that a distribution of best fit c parameters is needed for observed lensing clusters in order to provide a sample large enough to make an adequate comparison with simulations. A study of observed concentration values for clusters by Comerford & Natarajan (2007) show that the best fit lensing-derived c parameters are systematically higher than concentrations derived via X-ray measurements, a difference which can be at least

partly explained by effects of triaxiality of cluster halos (Corless & King 2007; Gavazzi 2005; Oguri et al. 2005; Clowe et al. 2004) or the substructure within the clusters (King & Corless 2007), although the latter effect may also produce a negative bias of c values. In addition, baryonic physics can increase the concentration parameter mildly by up to 10% as compared to dissipationless dark matter in pure dark matter simulations (see e.g. Lin et al. 2006).

5.5.3 Centre position

In addition to the three centre positions tested in Sect. 5.4, we computed κ maps with the inversion method of Kaiser & Squires (1993; KS93) and utilised the peak of this surface mass distribution as a fourth cluster centre. The KS93 method assumes that $\gamma = g$, which is not a good approximation near the centres of massive systems. Therefore, in comparison with the other methods, it provided us with a reference point as for how large variation one can reasonably expect for the various centroiding methods.

All centre positions obtained with the four methods lie within $1'$ and hence represent the errors expected when using the peak of a κ map as cluster centre. As can be seen from Table 5.1, varying the centre position only slightly can lead to different mass estimates. Although within error bars, the results from fitting NFW using a fixed c varies most. The NFW fitting of two parameters is more stable with a smaller spread in M_{200} . This is also reflected in χ^2/Dof , as a value closer to 1 is a better fit.

Worth noticing is the generally smaller differences between the results of Abell 1995 as compared to those of Abell 1351. The concentration parameter also seems better constrained for Abell 1995, where we could not obtain an upper limit for c only in the case where the BCG was used as the centre reference. On the other hand, an upper limit for c could not be obtained for Abell 1351 for any of the cluster centres chosen. This is consistent with the fact that Abell 1351 is not in dynamical equilibrium, lacking a well-defined cluster centre. The results obtained from fitting spherically symmetric models hence depend on the cluster centre chosen.

5.5.3a Strong lensing features

Strong lensing effects (Sects. 2.1 and 2.2.2) are in general susceptible to substructures in clusters, and might constrain the centre of mass further. For both clusters archival WFPC2 HST data exist, taken for an ongoing snapshot survey of X-ray luminous clusters (HST PID 11103, PI: H. Ebeling). The images are taken through the F606W filter totalling 1200s exposure time each.

Taking into account both the colours and morphologies of galaxies in our $V - I$ data and the morphologies in the HST images, there are at least half a dozen plausible arcs and arclets visible in each of the two clusters. The lensing pattern for Abell 1351 appears to be very complex and does not indicate a single, well-defined centre. This is supported by the presence of several elliptical galaxies which are of similar brightness as the BCG. On the contrary, for Abell 1995 several arc(let)s are well aligned around the BCG (apart from three which are obviously associated with individual cluster galaxies), justifying adopting the BCG as cluster centre for Abell 1995. Applying strong lensing to our data will therefore not offer further constraints on the determination of the centre of mass than we already have.

5.5.4 The mass reconstructions

In Sect. 5.3 we presented the weak lensing reconstruction of the clusters' surface mass density, detecting the clusters on the 5σ level. In case of Abell 1351 a significant neighbouring peak A was detected, which coincides with the galaxy distribution. A faint radio emission filament is also pointing towards this peak (Giacintucci et al. 2009), further supporting its authenticity.

We note significant differences comparing our mass maps to those of Dahle et al. (2002), who used the KS93 inversion method. Abell 1995 appears rather circular in both reconstructions, with attached filamentary structures of low significance seen in the KS93 map. In their re-analysis of the Abell 901 supercluster field, Heymans et al. (2008) argued that such filamentary structures could be enhanced by the KS93 method itself. This algorithm assumes $g = \gamma$ near the critical cluster cores, which hence overestimates κ , and the smoothing implemented could then lead to the apparent merging of neighbouring peaks.

Our map of Abell 1351 appears roughly circular at the position of the main cluster with a significant extension towards the northeast, whereas it shows an extension to the southwest in the work of Dahle et al. (2002). Both reconstructions have peak B in common. Since it is not associated with any overdensity of galaxies but appears in both reconstructions based on very different data sets, the most likely explanation is a chance alignment of galaxies triggering this detection.

5.6 Conclusions

Utilising observations from CFH12K we find the masses of Abell 1351 (Abell 1995) to be $M_{200} \equiv M(r_{200}) \sim (8-9) \times 10^{14} h_{70}^{-1} M_{\odot}$ ($M_{200} \sim (5-6) \times 10^{14} h_{70}^{-1} M_{\odot}$). These results are derived from fitting our data to the NFW profile, altering both c and r_{200} , to find the best theoretical fit to our shear measurements (Fig. 5.7). Our κ reconstructions yield mass estimates of $M_{2D}(<r_{200}) = (11.7 \pm 3.1) \times 10^{14} h_{70}^{-1} M_{\odot}$ ($M_{2D}(<r_{200}) = (10.5 \pm 2.7) \times 10^{14} h_{70}^{-1} M_{\odot}$) for Abell 1351 (Abell 1995).

Our results illustrate that using solely weak lensing measurements, with no photometric or spectroscopic information to separate cluster members from background galaxies, the concentration parameter found for a galaxy cluster is poorly constrained. Future spectroscopic redshift measurements and strong lens modelling of the arcs seen towards the cores of Abell 1351 and Abell 1995 might help improve the constraints on their concentration parameters. However, the change in c value resulting from small variations of the centre position of Abell 1351 indicates that the mass distribution in the core of this dynamically unrelaxed cluster is too complex to be meaningfully fit by the NFW mass density profile. M_{200} , on the other hand, varies only slightly for both clusters when either r_{\min} is decreased or the cluster centre varied. Although the KSB+ PSF correction leads to an underestimation of the cluster masses (with a maximum of 20% in our case), increasing the background galaxy ellipticities correspondingly still yields masses within error bars of our current estimates. We therefore conclude that the mass estimates presented are robust.

Lensing by high-redshift galaxy clusters using Lyman-break galaxies in the COSMOS survey

Measuring the tangential shear profile of a lensing cluster of galaxies and comparing this profile to theoretical predictions enables us to calculate the projected cluster mass without making any assumptions about its dynamical state (Chap. 2). Since they are cosmologically young objects, high-redshift ($z > 1$) galaxy clusters are often still forming and not fully virialised. Applying gravitational lensing to clusters of galaxies is thus especially useful in studies of the earlier Universe at a time when it was half its current size (Chaps. 1 and 3). The abundance and mass of $z > 1$ clusters are extremely sensitive to cosmological parameters, in particular Ω_m (the matter density of the Universe; Sect. 1.5) and its fluctuation σ_8 (e.g. Warren et al. 2006; Donahue et al. 1998). The mass properties of such clusters are therefore worth looking into in detail.

The background galaxy population becomes fainter with higher redshift, leading to a higher fraction of sources that are unresolved in ground-based data, thus requiring space-based data for reliable shear measurements. The Cosmic Evolution Survey (COSMOS; Scoville et al. 2007b,a) is excellent in this respect, providing deep space-based, high-resolution images from the Hubble Space Telescope (HST) for a large area on the sky (Schrabback et al. 2010; Massey et al. 2007b). Using space-based data free from atmospheric seeing therefore also reduces the systematic uncertainties in the shear measurements since smaller shape corrections are required.

Utilising COSMOS data, we study a set of 11 galaxy clusters in the redshift range $z = 1.2 - 1.6$ using weak lensing. Such studies require numerous background galaxies to be analysed. The clusters' high redshifts put further constraints on the source catalogue, as the galaxies must be located far behind the cluster in order to produce a lensing signal. The uncertainties of photometric redshift (photo- z ; Sect. 4.4.1) estimates increase with redshift. However, the Lyman-break technique provides a slightly different method much more reliable in choosing galaxies at very high redshifts ($z > 3$), so-called Lyman-break galaxies (LBGs, Sect. 6.1.4). These are objects whose spectral distribution shows specific features confirming their high-redshift nature (Sect. 6.1.4). However, the number densities of galaxies with redshifts $z > 2$ are very low, and their magnitudes are faint. This leads to

increased uncertainties in the shape measurement and correspondingly larger error bars. We therefore also include additional source galaxies in our cluster analysis.

Our shear catalogue can be divided into two sets; the first contains galaxies with assigned photometric redshifts (photo- z s), the second comprise those that were typically too faint for photo- z determination. LBGs are present in both. In this chapter we determine cluster masses using various combinations of these source catalogues, and compare their performances.

6.1 The data set

6.1.1 The Cosmic Evolution Survey

The COSMOS survey comprises imaging and spectroscopy across a wide wavelength range taken with different telescopes. It has been used for many scientific applications, such as galaxy evolution, star formation, active galactic nuclei, and large-scale structure (Scoville et al. 2007b,a).

The HST part of the survey was carried out with the Advanced Camera for Surveys (ACS). It is centred on RA = $10^{\text{h}}00^{\text{m}}28^{\text{s}}.6$, DEC = $+02^{\circ}12'21''.0$. With 590 orbits it is the largest HST survey ever made, covering an area of 1.64 deg^2 ($\sim 77' \times 77'$) with 50% completeness for sources with $d = 0''.5$ at $I_{\text{AB}} = 26$ magnitude (Sect. 4.3.1). The data consist of 579 ACS tiles, each observed in F814W for 2028 s, taken between October 2003 and November 2005. We employ the boundaries defined by Ilbert et al. (2009), $149^{\circ}411 < \text{RA} < 150^{\circ}827$, $1^{\circ}499 < \text{DEC} < 2^{\circ}913$, inside which all filters (Sect. 4.3.2) utilised in the photo- z estimation have a uniform and deep coverage (see also Sect. 6.1.5).

6.1.2 The cluster list

We use the list of high-redshift galaxy clusters in the COSMOS field presented by Zatloukal et al. (2007), hereafter Z07. This list contains 15 cluster candidates in the redshift range $1.22 \leq z \leq 1.55$, selected as excess densities in the 3-dimensional galaxy distribution compared to the average object density in the surrounding field. Only overdense objects more than 3σ above the average field were considered in the process. The method is described in detail in Röser et al. (2010).

Due to their low masses given by Z07 we discard two of the cluster candidates (#1 and #15). These mass estimates are calculated from the field-corrected total luminosity in the rest-frame V band, assuming a present-day mass-to-light ratio of 300 extrapolated to the cluster redshifts. Furthermore, two of the candidates lie outside the D2 field-of-view (#5 and #9), leaving 11 clusters for our study.

6.1.3 The shear catalogue

We use the catalogue of Schrabback et al. (2010; hereafter S10), which contains shapes for 451 384 galaxies with $i_{814} < 26.7$. In their paper, S10 present a comprehensive analysis of weak lensing by large-scale structure in the COSMOS survey to constrain cosmological parameters, in addition to studying the redshift dependence of the lensing signal. We refer the reader to S10 for specifics regarding the image reduction, PSF correction and shear extraction (see also Chap. 4 and Sect. 2.3).

6.1.4 The Lyman-break catalogue

High-redshift galaxies ($z \sim 3 - 5$) in our shear catalogue are identified using the Lyman-break technique (Steidel et al. 1996; Giavalisco 2002). Hildebrandt et al. (2009; hereafter H09) use this method to identify LBGs in the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) Deep fields, in which 1 of 4 fields, namely D2, covers the innermost 1 square degree of the COSMOS area.

Galaxies generally have a high abundance of hydrogen, which has a large ionisation cross-section. This leads to an effective absorption of photons with $\lambda < 912 \text{ \AA}$ by (neutral) hydrogen in the source. Furthermore, intergalactic hydrogen clouds between the source and observer absorb practically all photons with $\lambda \lesssim 912 \text{ \AA}$, in addition to a large fraction of those with $\lambda < 1216 \text{ \AA}$. In the spectrum of such LBGs this can be seen as a significant decrease of flux bluewards of 1216 \AA (the so-called Lyman-break). This means that galaxies at a redshift higher than $z \sim 2.5$ essentially have no flux in the u -band filter, and are thus called u -band dropouts. Starting from redshifts ~ 3.5 onwards, galaxies also will drop out of the g -band filter, and so on. Since the photo- z method does not include spectral energy distribution (SED) templates (see Sect. 4.4.1 for details) specifically made for LBGs, this method does not work particularly well for such galaxies. However, since only a small fraction of galaxies in a magnitude-limited sample are LBGs, the consequence is minor. The Lyman-break method, on the other hand, is designed to provide an extremely pure sample of high-redshift galaxies, though not necessarily complete.

The LBGs are selected according to Sect. 3.2 in H09. To summarise, regions of high detection efficiency and low contamination are identified in colour-space to select u -, g -, and r -dropouts. All dropouts are required to have a SExtractor CLASS_STAR parameter < 0.9 ; in addition, g -dropouts cannot be detected in the u band, and r -dropouts cannot be detected in the u nor in the g band. Following this selection scheme the LBG catalogue consists of 8053 u -, 13 043 g -, and 4040 r -dropouts, the dropouts being mutually exclusive.

By utilising a cut in magnitude of $r > 23.0$ for u -, $i > 23.5$ for g -, and $z > 24.0$ for r -dropouts, H09 ensure the contamination from stars and low- z interlopers is below 10%. Further improvements come from removing LBGs inside star and star halo masks, which are prone to be wrongly classified as LBGs due to scattered light from the neighbouring star. We then match the LBG catalogue and the shear catalogue to obtain a high-quality LBG shear catalogue consisting of 4704 u -, 5473 g -, and 1191 r -dropouts. Table 6.1 shows the number of LBGs in the original catalogue, in addition to the masked and unmasked versions with magnitude cut. The numbers for the matched LBG shear catalogues are also shown.

The redder the dropout filter, the fewer galaxies compared to the original number comprise shear measurements. We investigate this effect further by looking at the limiting magnitudes in the CFHT i band utilised by H09 for the masked samples (see Table 6.1), where i_{lim} is defined as the 50% completeness limit (the magnitude where the number of galaxies drops to 50% of the expected value). For the u -dropouts i_{lim} is the same before and after the matching between the LBG and the shear catalogues, which was also expected due to most of these galaxies having a shear counterpart. The difference in i_{lim} is 0.3 magnitude for both the g - and the r -dropouts, pushed slightly towards higher magnitudes for the reddest sample. The match between shear and LBGs goes down correspondingly. For the r -dropouts an additional reason arises. Since their spectrum is moving out of the

Table 6.1: Number of Lyman-break galaxies in samples of different selection criteria. The limiting magnitudes for the masked samples are given at the bottom of the table.

	#u-dr.	#g-dr.	#r-dr.	#all dr.
Original	8053	13 043	4040	25 136
Unmasked:				
Mag cut	7930	12 961	4014	24 905
Matched	5834	6565	1463	13 862
Masked:				
Mag cut ^a	6256	10 231	3040	19 527
Matched	4704	5473	1191	11 368
i_{lim}	u-dr.	g-dr.	r-dr.	
Mag cut	26.1	26.8	27.0	
Matched	26.1	26.5	26.7	

^aNote that these numbers differ from those given in Hildebrandt et al. (2009) due to the more conservative masking scheme required by their science application.

Table 6.2: Number of galaxies in the shear catalogue.

Shear catalogue^a	
Matched to photo- z	194 440
Unmatched to photo- z	256 944
Sum	451 384

^aSchrabback et al. (2010)

i band, in which the lensing detections are made, a large fraction of these galaxies are not visible in the HST image.

6.1.5 The photometric redshift catalogue

The publicly available catalogue from Ilbert et al. (2009; I09) contains accurate photometric redshift estimates computed with 30 broad-, medium-, and narrowband filters covering the ultraviolet, visible, near-IR, and mid-IR, and has been filtered for $i < 25$ (see I09 for further details). It contains 385 065 objects identified with photo- z s, of which 369 435 are galaxies. For faint and high-redshifted galaxies the redshift uncertainties are given as $\sigma_z \sim 0.06(1 + z)$.

Matching the photo- z catalogue to the shear catalogue we identify 194 440 common galaxies, leaving 256 944 galaxies without an individual redshift estimation (Table 6.2). See Sect. 6.4 for details regarding the modelling of the redshift distribution of these galaxies.

6.2 Lensing measurements

Given the high redshifts of the clusters and the comparably small survey area of COSMOS, the clusters analysed in this study are likely to be low-mass systems since massive clusters are rare (Warren et al. 2006; e.g.). In order to obtain a lensing detection we therefore create a stacked background galaxy catalogue containing sources within $r_{\max} = 2.0 h_{70}^{-1}$ Mpc around each cluster, utilising the centre coordinates given by Z07 (see Table 6.5). Before we can statistically combine the lensing signals of the clusters, we have to rescale their values to a common system. To this end we chose the mean of the cluster redshifts, $\langle z_{\text{cl}} \rangle = 1.357$, as the representative lens redshift, and we reproject the lensed sources to $z_s = 3.0$. Since the Einstein radius of a cluster depends on the redshift of the lensed source, this projection also ensures that results using different background catalogues can be directly compared. To suppress noise and benefit from galaxies with smaller measurement errors, we also apply weights to all galaxies. The scaling to a reference cluster and source redshift can be included in the weighting process.

A reduced shear estimator for each galaxy i , $E_i = g_i/n_i$ (Sect. 2.2.4), is computed from its corresponding reduced shear, g_i , and used when calculating the average tangential shear of the clusters. The ratio between the lens efficiency factor of the original cluster/source redshifts and that of the reference redshifts is given by (where the foreground source galaxies have been removed from the catalogue, see eq. (2.34))

$$n_i = \left(\frac{D_d D_{\text{ds},i}}{D_{\text{s},i}} \right)_{\text{cl}} / \left(\frac{D_d D_{\text{ds}}}{D_s} \right)_{\text{ref}} . \quad (6.1)$$

The individual weights are then defined by the measurement error of each galaxy,

$$w_i = \left(\frac{1}{\sigma_{E_i}} \right)^2 = \left(\frac{n_i}{\sigma_{g_i}} \right)^2 . \quad (6.2)$$

We fit the SIS profile (Sect. 2.4.2) to the tangential shear of the stacked background catalogue, including the weights, eq. (6.2), in the process. Given the fixed cluster centre, this model is left with one free parameter, the Einstein radius θ_E .

6.2.1 Inner cluster radius

Due to cluster galaxy contamination close to the cluster centre (Sect. 6.4.3), lensing measurements can only be made starting from an inner radial cutoff, r_{\min} , onwards. However, this restriction does not apply to our LBG catalogue, as the Lyman-break technique assures these galaxies are indeed behind the cluster in question (Sect. 6.3). On the other hand, shear measurements of the innermost parts of the cluster are made from very few background galaxies only, due to the small area in question. Given the steep profile of the SIS model at small radii, the outcome of the SIS fit to the data is highly dependent on the lensed galaxies close to the cluster centre. A small number of noisy galaxies at small radii can therefore strongly affect the whole fit. Stable results are therefore obtained by excluding the innermost galaxies and starting from a minimum radius, r_{\min} .

Figure 6.1 shows the resulting θ_E vs. r_{\min} for LBGs (left) and nonphoto-z galaxies (right). It is seen in both plots that the Einstein radius is increasing with starting radius

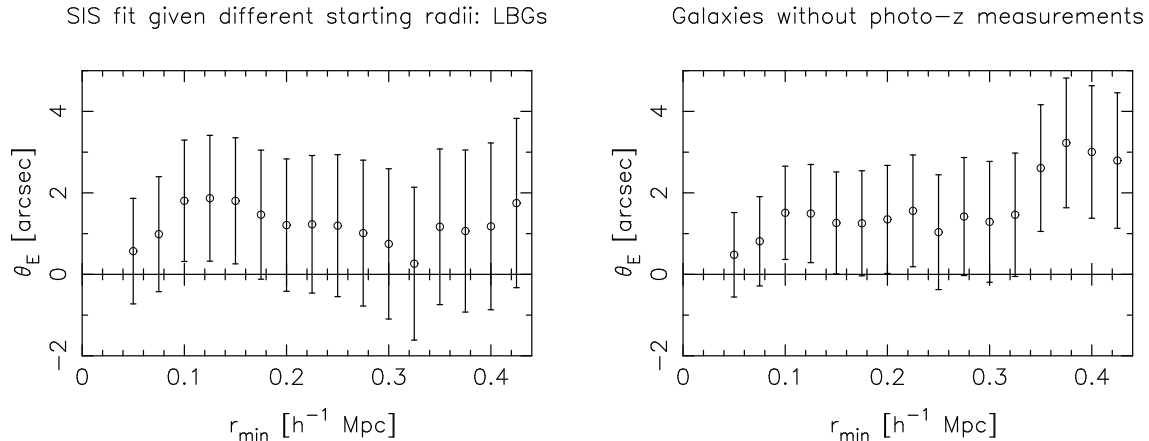


Figure 6.1: Resulting Einstein radius as a function of starting radius in the SIS fit. *Left*: Results for LBG catalogue alone. *Right*: Corresponding results for the background galaxy catalogue without photometric redshift information.

until $r_{\min} = 0.1 h_{70}^{-1}$ Mpc, after which the curve flattens out within the noise. To avoid biasing we therefore choose a starting radius $r_{\min} = 0.1 h_{70}^{-1}$ Mpc.

This starting radius also absorbs certain uncertainties with respect to the cluster centre position. To avoid centroiding issues (see the end of Sect. 7.4, page 114) we would ideally like to investigate the data ourselves and e.g. pick each cluster’s brightest cluster galaxy (BCG) as cluster centre to compare. However, the H -band data used by Z07, in which the red sequence galaxies (Sect. 3.1) at such redshifts are the brightest, are not publicly available. We therefore work with the cluster centres given.

6.3 Modelling the lensing signal from Lyman-break galaxies

LBGs have a naturally small dispersion in redshift, see Fig. 6.2, in addition to being so distant that the relative lensing strength in this redshift bin changes very little. There is therefore no need to calculate the redshift distribution of our LBG source population. Instead we utilise the median redshift for each dropout catalogue separately, calculated from a subset of the corresponding dropout sample where galaxies with assigned low redshifts have been excluded (where the photo- z estimation fails despite the clear dropout feature). The redshift ranges of these high-redshift catalogues are $2.17 \leq z_u \leq 3.77$, $3.33 \leq z_g \leq 4.50$, and $4.48 \leq z_r \leq 5.11$, with median redshifts of 3.25, 3.66, and 4.71, respectively. Note that all of the redshifts in this section are calculated by H09.

We create a stacked LBG catalogue following the approach in Sect. 6.2. The results utilising the u -, g -, and r -dropout catalogues separately are found in Table 6.3 and Fig. 6.3 (note that the data are binned for visualisation purposes only). The r -dropouts do not appear to contribute much signal. This is most likely a consequence of their faint magnitudes, making their shear measurements difficult, in addition to the comparatively small number of matched dropouts. For example, the mean HST i_{814} -band magnitude of the r -dropouts in this catalogue is 26.07 ± 0.43 , compared to $\langle i_{814} \rangle_u = 25.51 \pm 0.64$ for u -dropouts and $\langle i_{814} \rangle_g = 25.88 \pm 0.53$ for g -dropouts (where the error bars represent the intrinsic scattering of each sample). We therefore discard these galaxies from the final SIS fit. The

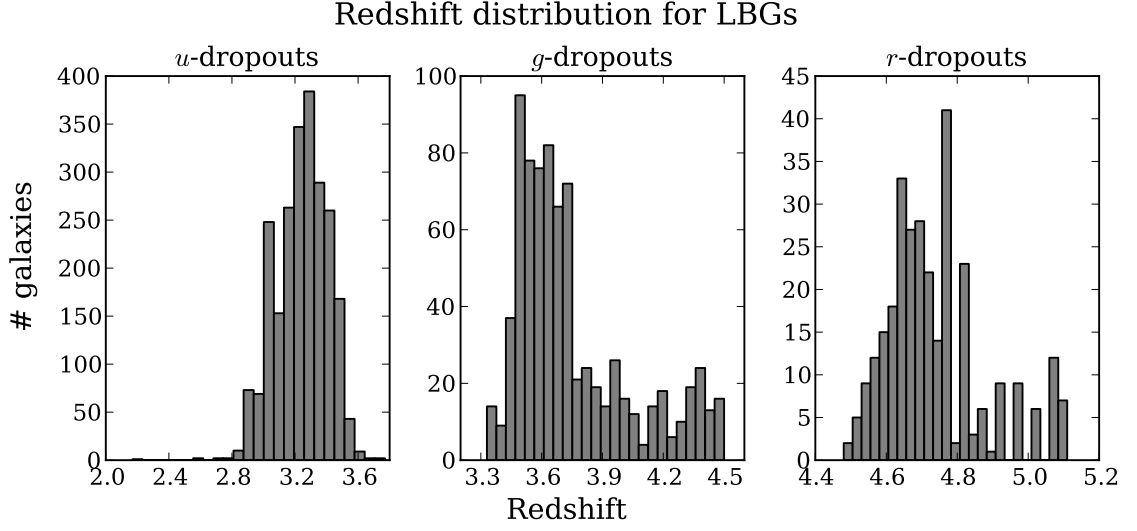


Figure 6.2: Redshift distribution for the Lyman-break galaxies (u -, g -, and r -dropouts). It is seen that all dropout samples have only a small dispersion in redshift. Dropout galaxies with wrongly assigned low photometric redshifts have been excluded from the figure.

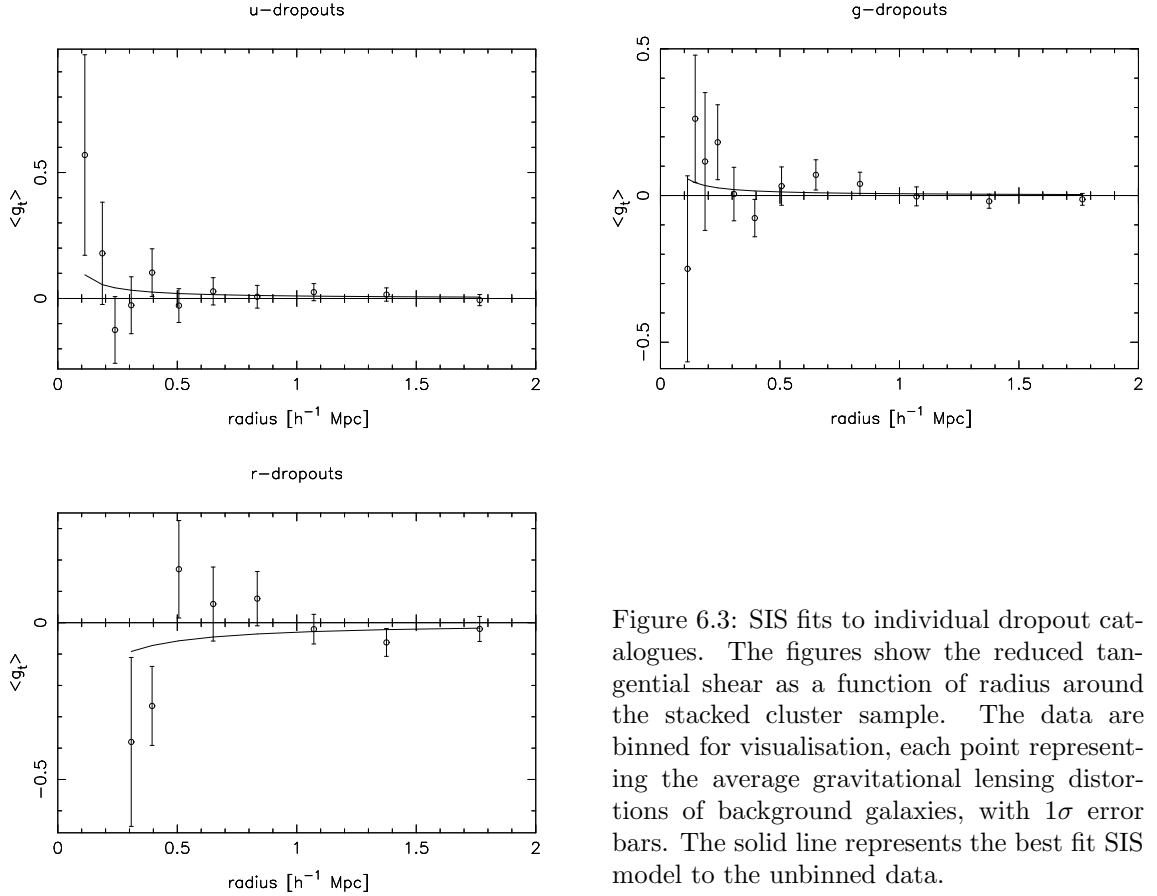


Figure 6.3: SIS fits to individual dropout catalogues. The figures show the reduced tangential shear as a function of radius around the stacked cluster sample. The data are binned for visualisation, each point representing the average gravitational lensing distortions of background galaxies, with 1σ error bars. The solid line represents the best fit SIS model to the unbinned data.

Table 6.3: Results from SIS fits to individual dropout catalogues.

dropout cat	θ_E ["]
u	2.31 ± 2.33
g	1.45 ± 1.94
r	-7.38 ± 4.44
u, g	1.81 ± 1.49
u, g, r	0.99 ± 1.41

 Table 6.4: Results from fitting the SIS profile to different background galaxy samples. Masses are calculated for $r = 1.0 h_{70}^{-1}$ Mpc. Clusters are scaled to $\langle z_{cl} \rangle = 1.36$ and source galaxies to $z_s = 3.0$. Note that the LBG sample includes u - and g -band dropouts.

Galaxy sample	θ_E ["]	M_{SIS} [$10^{13} h_{70}^{-1} M_{\odot}$]
LBG	1.81 ± 1.49	8.2 ± 6.7
photo-z	0.76 ± 1.48	3.5 ± 6.7
nonphoto-z	1.51 ± 1.15	6.8 ± 5.2
non- & photo-z	0.42 ± 0.28	1.9 ± 1.3
photo-z & LBG	1.81 ± 1.05	8.2 ± 4.8
nonphoto-z & LBG	2.15 ± 1.44	9.7 ± 6.5
final sample	1.98 ± 1.03	9.0 ± 4.7

fit to the remaining source catalogue (u - and g -dropouts; hereafter LBG catalogue) yields $\theta_E = 1''.81 \pm 1''.49$, corresponding to a typical mass of $M_{\text{SIS}} = (8.2 \pm 6.7) \times 10^{13} h_{70}^{-1} M_{\odot}$ within $r = 1.0 h_{70}^{-1}$ Mpc. The result is shown in Fig. 6.4, top panel. In this figure the left plots show the reduced tangential shear as a function of radius, and the right plots its cross-component (see Sect. 6.6.1 for a discussion regarding these B-mode measurements).

6.4 Modelling the lensing signal from non- and photo- z galaxies

Ideally one would like all shear galaxies to have redshift information. An easy way to reduce suppression of the shear signal from cluster members and (within errors) make sure only background galaxies are chosen, is to utilise only photo- z galaxies. Applying the method described in Sect. 6.2, taking individual galaxy redshifts into account, we fit θ_E to the sample of background shear galaxies containing photo- z s, choosing only galaxies with $z_s \geq (z_{cl} + 0.2)$ in the process (Abate et al. 2009). Note that LBGs are present in both the photo- z and the nonphoto- z catalogues. When analysing non- and photo- z galaxies separately or together, LBGs are not identified nor excluded, i.e. these galaxies remain in their respective catalogues and are treated correspondingly. The result yields $\theta_E = 0''.76 \pm 1''.48$, see Table 6.4. This is smaller than, but still consistent with what has been found with the LBG catalogue. The second panel of Fig. 6.4 shows the shear profile.

6.4.1 Modelling the redshift distribution for shear galaxies lacking photo-z estimates

Roughly 40% of the galaxies in the shear catalogue are matched with objects in the photo-z catalogue (Table 6.2). In order to also use the remaining galaxies, we model their redshift distribution according to S10. Note that the procedure described below applies to the faint nonphoto-z galaxies only, i.e. no photo-z galaxies are included in the final redshift distribution (they are only used for comparison reasons).

The mean redshift of the photo-z galaxies within the magnitude range $23 < i_{814} < 25$ is given by

$$\langle z \rangle = (0.276 \pm 0.003)(i_{814} - 23) + 0.762 \pm 0.003 . \quad (6.3)$$

This relation also agrees very well with galaxies over the extended magnitude range $23 < i_{814} < 27$ in the Hubble Deep Field-North (Fernández-Soto et al. 1999; HDF-N), on average to 2%, for which redshifts exist for the complete catalogue. In comparison, the mean photo-zs for galaxies in the Hubble Ultra Deep Field (Coe et al. 2006; HUDF) are on average higher than eq. (6.3) by 16% for $23 < i_{814} < 25$ and 10% for $25 < i_{814} < 27$. The sampling variance owing to the small COSMOS field size can hence be roughly estimated from the difference between the HDF-N and HUDF. Since the HDF-N agrees very well with eq. (6.3), we can use this equation to estimate the mean redshift of our faint nonphoto-z sample (for which $i_{814} < 26.7$, see Sect. 6.1.3).

In addition to the mean redshift, the redshift distribution of the galaxies must also be determined, since the shear signal has a non-linear dependence on redshift. This redshift distribution can be parametrised as

$$p(z|i_{814}) \propto \left(\frac{z}{z_0}\right)^\alpha \left(\exp \left[- \left(\frac{z}{z_0}\right)^\beta \right] + cu^d \exp \left[- \left(\frac{z}{z_0}\right)^\gamma \right] \right) , \quad (6.4)$$

where $z_0 = z_0(i_{814})$ and $u = \max[0, (i_{814} - 23)]$, and $(\alpha, \beta, c, d, \gamma) = (0.678, 5.606, 0.581, 1.851, 1.464)$ are the best-fit parameters found by S10 (see their Sect. 2.2.2). Combining the two equations above, S10 estimate the fitting formulae

$$z_0 = 0.446(i_{814} - 23) + 1.235 \quad \text{for } 22 < i_{814} \leq 23 , \quad (6.5)$$

$$z_0 = \sum_{j=0}^{j=7} a_j \left[\frac{(i_{814} - 23)}{4} \right]^j \quad \text{for } 23 < i_{814} < 27 , \quad (6.6)$$

for which $(a_0, \dots, a_7) = (1.237, 1.691, -12.167, 43.591, -76.076, 72.567, -35.959, 7.289)$. The total redshift distribution then reads

$$\phi(z) = \sum_{k=1}^N \frac{p(z|i_{814})}{N} , \quad (6.7)$$

yielding a good description of the magnitude-dependent exponential redshift tail, as well as the peak of the redshift distribution.

Once the redshift distribution is determined,

$$\langle \beta \rangle = \left\langle \frac{D_{\text{ds}}}{D_{\text{s}}} \right\rangle \quad (6.8)$$

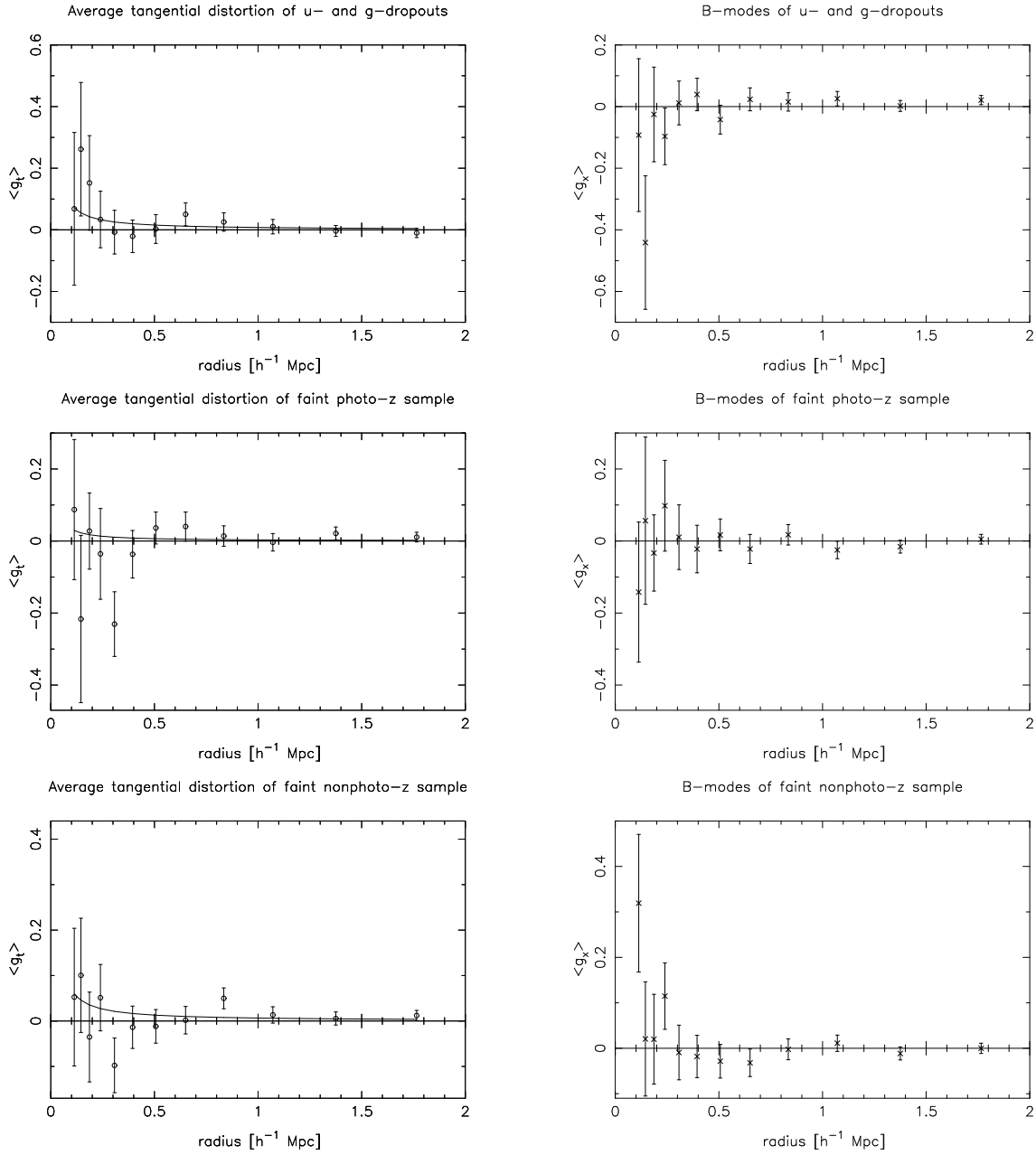


Figure 6.4: SIS fit to (*from top to bottom*): LBG catalogue, background galaxies with photometric redshift information, background galaxies without individual redshift information, and final (combination of LBGs, photo- z , and nonphoto- z galaxies) background catalogue (*next page*). *Left*: Reduced tangential shear as a function of radius around the stacked cluster sample. The data are binned for visualisation, each point representing the average gravitational lensing distortions of background galaxies, with 1σ error bars. The solid line represents the best fit SIS model. *Right*: Cross-component of the reduced tangential shear. Such B modes can be used to test for systematics still present in the data, and should disappear if $\langle g_t \rangle$ is caused by lensing only. The figure shows that all B-mode profiles meet this condition (see Sect. 6.6.1 for more details).

6.4. Modelling the lensing signal from non- and photo-z galaxies

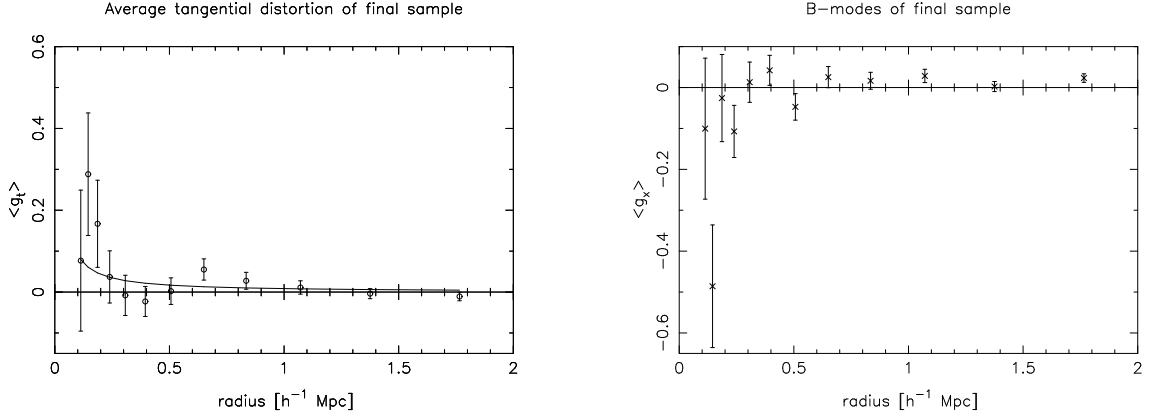


Figure 6.4 (continued)

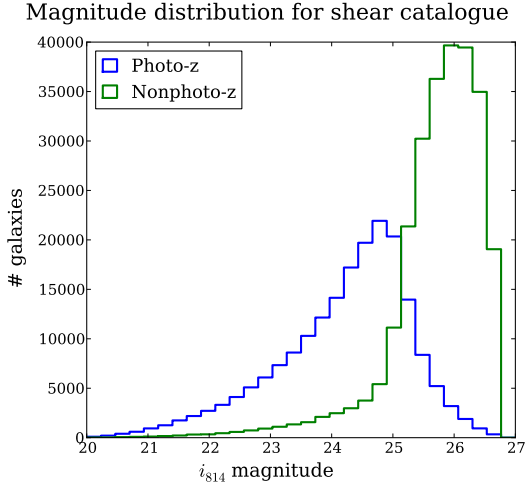


Figure 6.5: Magnitude distribution for the shear catalogue. The galaxies with photo-zs are drawn in blue (low peak), and those without redshift information in green (high peak). The i_{814} magnitudes originate from the space-based shear catalogue, and thus deviate from those of the ground-based photo-z catalogue, which has been filtered for $i < 25$. Note that LBGs have not been identified nor removed from the samples.

can be calculated for each cluster upon stacking, enabling us to fit the SIS model to galaxies both with and without photo-zs. The results can be found in Table 6.4.

6.4.2 Photo-z vs. nonphoto-z galaxies

As shown in Table 6.4, the SIS fit to galaxies with individual photo-zs did not yield a robust detection. On the other hand, for the sample of galaxies without individually known photo-zs we do obtain a positive detection of $\theta_E = 1''.51 \pm 1''.15$. The shear profile for this fit is seen in the third panel of Fig. 6.4.

We therefore investigate the mean i_{814} -band magnitudes of the two samples in more detail, and find that $\langle i_{814} \rangle_{\text{phz}} = 24.20 \pm 1.09$ for the photo-z galaxies and $\langle i_{814} \rangle_{\text{noz}} = 25.71 \pm 0.78$ for the nonphoto-z galaxies (where the errors indicate the intrinsic scattering of the respective samples). Going ~ 1.5 magnitudes deeper, the nonphoto-z galaxies are on average farther away than the photo-z sample. This is also seen in Fig. 6.5, where the i_{814} magnitudes of all galaxies are plotted in a histogram. However, examining the lensing efficiency (Sect. 2.2.4) for the two catalogues we find that $\langle \beta \rangle_{\text{noz}} = 0.11$ for the nonphoto-z sample and $\langle \beta \rangle_{\text{phz}} = 0.24$ for the photo-z sample. The dilution from the large amount of foreground objects in the nonphoto-z catalogue is hence bringing down the

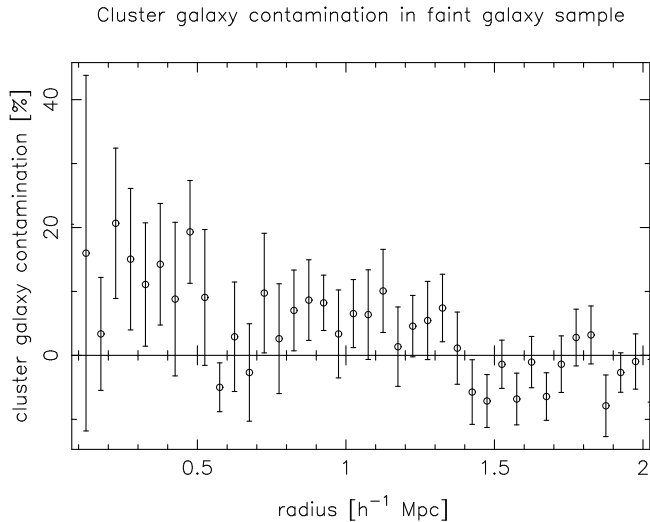


Figure 6.6: Cluster galaxy contamination in nonphoto- z shear catalogue for stacked cluster sample. Each value corresponds to the contamination within one radial bin and is plotted at the average radius for that bin. Data points are mutually independent.

lensing efficiency in this sample significantly. Nevertheless, since the lensing signal from the nonphoto- z galaxies produces a detection of higher significance than that of the photo- z sample, the faint galaxies must yield a remarkably high signal in order to dominate that of the total nonphoto- z catalogue.

6.4.3 Cluster galaxy contamination

Close to the centre of a galaxy cluster, the background galaxy catalogue is likely to be contaminated with cluster galaxies. For galaxies with photo- z measurements this effect is expected to be small, but for background galaxies without individual redshift information we need to quantify this contamination. We measure the overdensity of galaxies in radial bins around each cluster centre with respect to the corresponding field value. The mean of these values within each bin then gives the cluster galaxy contamination present in the stacked cluster sample as a function of distance from the cluster centre, see Fig. 6.6 (where the error bars are the errors of the mean). As seen in the figure, we detect a slight contamination at small scales ($\lesssim 15\%$). However, its impact is negligible compared to the statistical errors in the shear measurements, and we do not consider this effect further.

6.5 Fitting the SIS profile to the joint LBG, non- and photo- z catalogue

To optimise our weak lensing measurement we work with a combined shear catalogue containing the u - and g -band dropouts, in addition to the non- and photo- z samples defined in the previous section. The non- and photo- z shear catalogues are mutually exclusive, and together they comprise the total shear catalogue. The dropout galaxies are therefore excluded from these two catalogues to ensure no galaxy shear measurement is utilised twice.

The result from fitting the SIS profile to this final shear catalogue gives $\theta_E = 1''.98 \pm 1''.03$, which corresponds to a mass of $M_{\text{SIS}} = (9.0 \pm 4.7) \times 10^{13} h_{70}^{-1} M_{\odot}$ within $r = 1.0 h_{70}^{-1} \text{Mpc}$. The last panels of Fig. 6.4 display the fit (left) and the B modes (right).

6.6 Consistency checks

6.6.1 B modes

The right panel of Fig. 6.4 shows the cross-component of the reduced tangential shear as a function of radius. It corresponds to the curl of the surface mass density, $\nabla \times \kappa$, and is calculated by rotating all galaxies 45° and performing the same shear analysis as was applied to the original shear catalogue (Crittenden et al. 2002).

Such B-mode measurements can be used to test for residual systematics in the data, and should be consistent with Gaussian noise if $\langle g_t \rangle$ is caused by lensing only. This null test is consistent with zero for all shear measurements in Fig. 6.4, demonstrating that our analysis is largely free from systematics.

6.6.2 Weights

We made a sanity check of the weights by fitting the SIS profile to each separate and combined background galaxy catalogue without applying weights. In most cases these results are close to those obtained from the weighted catalogues, however, for the final galaxy sample as well as the one where nonphoto-z and LBGs are combined, θ_E is systematically twice as large although the detection significance remains the same. As the weighted results all agree within error bars (Table 6.4), the weights are clearly working well in suppressing noise in the faint, high-redshift part of the background galaxies.

6.6.3 Cluster redshift re-calculation

The cluster redshifts presented in Z07 are calculated independently of the I09 redshift catalogue used in this paper. Since the Z07 cluster detection method depends strongly on the redshift of each object (Sect. 6.1.2), we test its compatibility with I09 by calculating new cluster redshifts, utilising the median of all galaxies with $z_{cl} - 0.1 < z < z_{cl} + 0.1$ within a radius of $r = 1.0 h_{70}^{-1}$ Mpc around each cluster centre. The results are shown in Table 6.5. Due to masking in the photo-z catalogue, there are not enough objects present around the centres of clusters #3, #4, and #5 to calculate a robust median redshift (the projected sky coverage is less than $\sim 80\%$, see Sect. 7.2.4 for details). Since the photo-z catalogue is only utilised with respect to redshifts (the cluster richness/luminosity is not considered in this analysis), and the shear catalogue is complete around all cluster centres, we do not reject these clusters but use their original redshift estimate instead.

Utilising the re-calculated cluster redshifts yields $\theta_E = 2''.13 \pm 1''.42$ for the final background galaxy catalogue. This corresponds to a mass of $M_{SIS} = (9.6 \pm 6.4) \times 10^{13} h_{70}^{-1} M_\odot$ within $r = 1.0 h_{70}^{-1}$ Mpc, very close to the previous mass estimate of $M_{SIS} = (9.0 \pm 4.7) \times 10^{13} h_{70}^{-1} M_\odot$.

6.7 Conclusions

For the first time we measure the lensing signal from high-redshift galaxy clusters using higher-redshift Lyman-break galaxies. We show that even though these galaxies are efficiently lensed, the number density of our LBG catalogue is not high enough to sufficiently beat down the noise, and the resulting mass estimates carry large error bars.

Table 6.5: New calculation of cluster redshifts. Coordinates are taken from Z07.

ID	RA [°]	DEC [°]	Z07 ^a z	median z
1	150.148	2.061	1.22	1.25
2	150.586	2.093	1.23	1.19
3	150.551	1.816	1.24	-
4	150.599	1.758	1.24	-
5	150.227	2.784	1.26	-
6	150.591	2.179	1.31	1.27
7	150.100	2.699	1.32	1.29
8	149.991	2.690	1.33	1.28
9	150.263	2.763	1.40	1.43
10	149.882	2.608	1.41	1.44
11	149.843	2.115	1.43	1.45
12	149.960	2.337	1.44	1.45
13	149.976	2.490	1.45	1.45
14	149.910	2.328	1.52	1.48
15	149.775	2.457	1.55	1.57

^aZatloukal et al. (2007).

Weak lensing has been applied to high-redshift clusters before. Jee et al. (2009) study a massive cluster at $z = 1.4$, detecting it on an $\gtrsim 8\sigma$ level. This is consistent with our low lensing detection, given that the typical mass of a cluster in our sample is a factor of 10 smaller than that of Jee et al. (2009). Contrary to the cluster identified by Jee et al. (2009), the galaxy clusters studied in this chapter cannot be used to constrain cosmology due to the low signal-to-noise and intrinsically low mass.

To improve on our analysis further, we would have to reduce the noisy contribution from unlensed foreground objects. An easy solution would be to exclude all objects lacking redshift information. We have demonstrated that for high-redshift clusters the application of a magnitude limited ($i < 25$) photo- z catalogue does not yield the most significant detection. Most galaxies with $i > 25$ have a higher relative lensing efficiency than those in the photo- z catalogue. Since none (excluding LBGs in this comparison) of these galaxies have redshift estimates, the nonphoto- z sample yields a lensing detection of higher significance than that of the photo- z galaxies. This detection, however, is limited by the large contamination from foreground objects. We have consequently shown that the detection significance of a weak lensing analysis of high-redshift galaxy clusters depends on the limiting magnitude of the photo- z catalogue. If all, or a large fraction of the faint galaxies had redshift estimates, there might not be a need to include those without redshift information. Nevertheless, the redshift distribution estimation made by S10 proves to work well, demonstrating that despite lacking individual redshift information it is crucial for the results that these galaxies are properly included in the analysis.

Analysis of the mass-to-light ratio for galaxy clusters in the COSMOS field

7.1 Introduction

The comparison between luminous and total mass came to attention early in the 20th century, starting with Kapteyn (1922) who obtained values for the mass-to-light ratio of galaxies of approximately twice the value in the vicinity of the Sun. Hubble (1929c) found similar values for the Andromeda disk, whereas Zwicky was the first person to look at larger scales. As opposed to the previous studies, which found no dark matter present in galaxies, Zwicky concluded from his studies of the Coma cluster a significant dominance of dark (90%) over luminous (10%) matter (Zwicky 1933, 1937). These detections of dark matter in galaxy clusters became one of the pioneering results regarding the existence of dark matter (Sect. 1.8.1).

The cold dark matter (CDM; Sect. 1.8) model is currently the simplest model capable of accounting for the dominance of dark matter which has been observed over the last decades. Modern simulations are able to predict the evolution of structure in a CDM universe, where about 80% of the cosmic mass density is assumed to be in the form of some dynamically cold, collisionless species of the subatomic particle zoo. In contrast, the baryonic components of galaxy clusters are much more difficult to simulate due to the complex magneto-hydrodynamic processes involved. Observations are therefore critical for a deeper apprehension of the connection between the dark and luminous matter. The mass-to-light (M/L) ratio of groups and clusters is at the centre of attention in this respect, as they are cosmologically young objects undergoing strong evolution from $z \sim 1$ until today.

The M/L ratio gives the total amount of mass relative to the (optical) luminosity (Sect. 4.3.1), and specifies the relative contribution of the dark matter component (e.g. Girardi et al. 2002). We know from pioneering analyses that M/L increases with mass for cosmic objects in the Universe, from the bright luminous parts of galaxies out to cluster scales (e.g. Blumenthal et al. 1984). At cluster scales, however, M/L seems to approach a saturation value where it is not increasing with mass anymore, a feature which has also been detected when going to larger scales (for example superclusters; see e.g. Schneider 2006c). Although this conclusion has been made by several authors, other studies find

M/L to increase with mass from groups to massive clusters, as summarised in the following paragraph. A clear consensus is yet to be found.

Dressler (1978) analysed 12 clusters based on homogeneous optical data, revealing no evidence of correlation between M/L and richness. David et al. (1995) demonstrated a comparable M/L_V for 7 groups and clusters utilising X-ray masses and optical luminosity estimates (the latter obtained from literature). Using optical virial analyses (Sect. 3.3.1) to estimate the mass of 15 clusters from the Canadian Network for Observational Cosmology, Carlberg et al. (1996) also concluded with M/L_r values corresponding to an average M/L in the Universe. Hradecky et al. (2000) found that M/L_V and mass are roughly independent, after looking at 8 galaxy groups and clusters with homogeneous X-ray masses and optical luminosities. Kochanek et al. (2003) were looking at clusters both in N -body simulations and K -band data from the Two Micron All Sky Survey, also finding that M/L_K for the virialised region of clusters is roughly independent of cluster mass. Finally Sheldon et al. (2009) found that M/L is approaching a richness-independent value when going from small to large scales, this value being consistent with M/L within r_{200} (see Sect. 2.4.3) for large clusters.

There are, however, several papers coming to different conclusions. Studies of the cluster fundamental plane, i.e. the connection between the luminosity, velocity dispersion, and size of a cluster, show slight tendencies of an increasing M/L with mass (e.g. Schaeffer et al. 1993; Adami et al. 1998a,b). Other studies also conclude with a slight but significant tendency that the mass is increasing faster than the luminosity (e.g. Girardi et al. 2000; Carlberg et al. 2001; Girardi et al. 2002). Covering a large mass range in both optical and near-infrared bands, most authors find the relation between M/L and mass to be of the form $M/L \propto M^\alpha$, with $\alpha = 0.2 - 0.4$ (e.g. Bahcall & Comerford 2002; Lin et al. 2003, 2004; Rines et al. 2004; Ramella et al. 2004; Muzzin et al. 2007; the last paper finding $\alpha = 0.57 \pm 0.13$).

While the question regarding whether or not M/L of clusters are representative for the Universe as a whole is still unanswered, we assume in the following that it is, and proceed to calculate the matter density parameter of the Universe, Ω_m . Oort's method (Oort 1958; see also Carlberg et al. 1999) provides a direct measure of Ω_m , in that the mass density of a field can be calculated from the product between the field luminosity density and the total M/L . By letting ρ_L define the typical luminosity density of the Universe, usually calculated from field galaxies, Ω_m can be estimated from

$$\Omega_m = \frac{M}{L} \frac{\rho_L}{\rho_{\text{cr}}}. \quad (7.1)$$

Here ρ_{cr} is the critical density of the Universe (see eq. (1.31), Sect. 1.5), and M/L describes the total mass-to-light ratio of the Universe. Both ρ_L and L must be measured in the same band or frequency interval.

The above equation assumes that galaxy formation does not change between environments and that the luminosity of field galaxies remains the same when they fall into a cluster. However, both of these assumptions are questionable, which can also be seen from the discussion above, regarding the validity of assuming that cluster M/L ratios represent the universal $\langle M/L \rangle$. Studies of galaxy populations both inside and outside cluster environments also reveal that field galaxies differ from cluster galaxies (see e.g. Dressler 1984a,b). Typical results yield $\Omega_m \sim 0.15$ (e.g. Reiprich 2001), much lower than

the commonly assumed $\Omega_m \sim 0.3$.

In this chapter we continue the analysis of galaxy clusters in the Cosmic Evolution Survey (COSMOS; see Chap. 6), this time including clusters over a wide redshift range, $0.2 \leq z \leq 1.6$. We investigate the relations between their mass, luminosity, and richness in detail, as well as looking into how these relations evolve with redshift. In addition, we examine M/L as a function of mass. To avoid the correlation between M/L and mass, we also study the relation between M/L and cluster X-ray temperature (which is used as a proxy for mass).

7.2 Cluster samples

The clusters analysed in this chapter are taken from two different cluster lists, obtained using two very different cluster detection methods. The first list is derived by Zatloukal (2008; hereafter Z08) using optical detection methods, the second list by Finoguenov et al. (2007; hereafter F07) using X-ray analysis.

7.2.1 Optically selected clusters

The HIROCS optical cluster list of Z08 is based on a modified friends-of-friends (FoF) algorithm (e.g. Huchra & Geller 1982) running on the three-dimensional space spanned by sky coordinates and photometric redshifts (photo-zs; Sect. 4.4.1). The detection algorithm is summarised in the following.

For each galaxy position the local galaxy density is calculated, and the sample sorted according to density from high to low, counting only those above a defined overdensity cut. Taking the position with the highest overdensity as starting point, where the galaxy at this location has redshift z , a subsample is created in which all galaxies must lie inside a redshift slice of $z \pm 0.1$ and within a projected radius of 300 kpc. The 3 galaxies (including the starting galaxy) lying at the most overdense positions in this subsample are then chosen, their mean redshift forming the redshift of the cluster candidate.

Using the full galaxy sample, more cluster members are now identified within $\pm 2\sigma_z$ of the cluster redshift. Applying the FoF algorithm, all connected overdense objects within a 300 kpc search radius between galaxies are detected. These member galaxies are then removed from the full sample, and the procedure repeated with the remaining galaxies. Once all galaxies have been identified, a final cluster list is created, keeping only structures consisting of at least 6 galaxies and having a minimum overdensity of 3σ as compared to the field value.

Z08 construct their own photo-zs from $u, g, r, i, z, B, V, H, K$ broad-band filters and the *NB816* narrow-band filter. They establish two cluster lists: a low-redshift sample with $0.5 \leq z < 1.0$, and a high redshift sample with $1.0 \leq z < 1.6$. The low-redshift clusters extend over the full COSMOS field of 2 square degrees, whereas the high-redshift ones require deep coverage in the *H* band, which exists only for 0.66 square degrees.

The total cluster list consists of 172 clusters, their RA, DEC, and redshifts can be found in Table A.1.

7.2.2 X-ray selected clusters

The X-ray selected cluster catalogue is produced by F07 and contains 72 galaxy clusters. The clusters have been detected in the 0.5 – 2 keV band using *XMM-Newton* observations of the COSMOS field covering the entire 2 deg². In addition to standard data processing, a refined background subtraction is performed, creating a final X-ray image of fairly uniform signal-to-noise ratio (F07).

The clusters are chosen from the spatial extent of their X-ray emission using a two-step detection algorithm. Firstly areas with detectable flux on large angular scales are selected. Secondly those whose flux is dominated by the flux from point-like sources (mostly being active galactic nuclei) are removed (80% of the initially selected areas). Finally the X-ray cluster candidates are cross-correlated with a list of galaxy groups constructed from photo- z s, keeping only the clusters with a corresponding optical counterpart. F07 use the photo- z catalogue of Mobasher et al. (2007) based on i -band detections, containing galaxies with $i_{\text{AB}} < 25$ and having a 1σ redshift estimate uncertainty of $0.027(1+z)$. The cluster list is hence limited to $z < 1.3$, after which the 4000Å break moves redward of the i -band filter.

Galaxy overdensities are identified in redshift slices in photo- z space using only high-quality photo- z early-type galaxies (Sect. 3.1) that are not morphologically classified as stellar objects. The diffuse X-ray structure is identified as a cluster if it contains a galaxy density peak, and the cluster centre and redshift chosen from its strongest galaxy peak. Since each cluster is detected in several redshift slices, the strongest peak is considered representative for the most likely cluster redshift. In this study we select clusters with $z \geq 0.2$, resulting in a total of 65 X-ray clusters. Table A.2 lists the clusters identified this way by F07.

Most of the cluster flux is encompassed within r_{500} (see e.g. the study of nearby galaxy clusters by Markevitch 1998; r_{500} denotes the radius inside which the matter density equals 500 times that of the critical, see also Sect. 2.4.3). The total flux of the X-ray clusters is therefore measured within this radius. Following the approach by Böhringer et al. (2004), F07 calculate the rest-frame luminosity by iteratively taking the total flux within an estimated r_{500} and applying the K correction (Sect 7.3.1) accounting for the cluster’s redshift and temperature.

The relations between the X-ray luminosity ($L_{0.1-2.4\text{keV}}$) and the X-ray temperature (T_{X}) derived by Markevitch (1998) is used to estimate T_{X} of each cluster,

$$k_{\text{B}}T_{\text{X}} = 6 \text{ keV} \left(\frac{L_{0.1-2.4\text{keV}} h_{70}^2}{3 \times 10^{44} E_z \text{ ergs/s}} \right)^{0.48}, \quad (7.2)$$

where

$$E_z = \sqrt{\Omega_{\text{m}}(1+z)^3 + \Omega_{\Lambda}} \quad (7.3)$$

and k_{B} denotes the Boltzmann constant. Re-deriving the $M-T_{\text{X}}$ relation from Finoguenov et al. (2001) using orthogonal regression, the total gravitational mass and corresponding r_{500} are estimated through

$$M_{500} \equiv M(r_{500}) = 2.36 \times 10^{13} h_{70}^{-1} M_{\odot} \times \left(\frac{k_{\text{B}}T_{\text{X}}}{1 \text{ keV}} \right)^{1.89} E_z^{-1} \quad (7.4)$$

and

$$r_{500} = 0.391 h_{70}^{-1} \text{ Mpc} \times \left(\frac{k_{\text{B}} T_{\text{X}}}{1 \text{ keV}} \right)^{0.63} E_z^{-1}, \quad (7.5)$$

where the masses have been corrected to h_{70} and a Λ CDM cosmology.

7.2.3 Common clusters

Clusters are defined to have an X-ray counterpart if they fulfil these criteria:

- The X-ray centre lies within 1 Mpc from the optical cluster centre.
- The redshifts of the two clusters coincide within 2σ , fulfilling:

$$|z_{\text{xray}} - z_{\text{opt}}| \leq 2\sqrt{\sigma_{z,\text{xray}}^2 + \sigma_{z,\text{opt}}^2} \quad (7.6)$$

where z_{xray} and z_{opt} are the cluster redshift estimates of the X-ray and the optically selected clusters, respectively. Further, $\sigma_{z,\text{xray}} = 0.027(1+z)$ is the 1σ uncertainty of the redshift estimate for galaxies with $i_{\text{AB}} < 25$ in the X-ray catalogue (Mobasher et al. 2007), here adopted as the error of the cluster redshift itself, and $\sigma_{z,\text{opt}}$ is the error of the cluster redshifts estimated by the optical cluster selection method of Z08.

There are 16 clusters fulfilling this criterion, all marked under ‘‘Common’’ in Tables A.1 and A.2. For these clusters we then replace the initial positions of the cluster centres given by Z08 with the corresponding X-ray centres from F07. Given the spatial relocation and the corresponding change of cluster galaxies, we also adjust the redshifts to match those of the X-ray cluster catalogue.

It is peculiar that only 16 common clusters are found in both catalogues. The reason why so few of the optically selected clusters are detected in the X-ray has likely got to do with cluster masses – the optical clusters are simply not massive enough, i.e. they have either no or very faint X-ray emission. The question remains why not all X-ray clusters are found in the optical catalogue. Since the X-ray detections are confirmed using overdensities in the galaxy distribution, it is not immediately clear why they are not also found in the optical cluster search. Possible reasons include issues regarding the photo- z accuracy, for which Mobasher et al. (2007) use 16 bands and Z08 only 10, or problems finding the correct cluster centre (see the end of Sect. 7.4, page 114, for a discussion regarding this topic).

7.2.4 Projected sky coverage

In this study we use the COSMOS field boundaries (Sect. 6.1.1) as defined by Ilbert et al. (2009; hereafter I09). Because Z08 include private data and compute their own redshift measurements, 6 of their clusters are outside this area. The boundaries also exclude one of the X-ray clusters.

Due to (star and other) masks (Sect. 4.2.2), the projected sky coverage of the photo- z and shear catalogues must be evaluated. A cutout with projected radius $r = 1 \text{ Mpc}$ ($r = 2 \text{ Mpc}$) around each cluster centre is created from the photo- z (shear) catalogue, and

Table 7.1: Number of clusters in the optical and X-ray cluster lists, as well as in the different redshift bins. Note that the optical and X-ray cluster classifications are mutually exclusive. When common clusters are introduced to the list all three classifications are mutually exclusive. In the fitting procedures the common clusters are included in both the optical and the X-ray samples.

	Optical	X-ray	Common
Total	172	65 ^a	
Rejected due to being outside I09 ^b field ^c	6	1	
Rejected due to low shear coverage	25	6	
Rejected due to low photo-z coverage	27	9	
Total used in this study	127	51	
Total used in this study	114	38	13
Redshift bins:			
$0.2 \leq z < 0.4$	0	13	0
$0.4 \leq z < 0.6$	14	7	2
$0.6 \leq z < 0.8$	34	3	3
$0.8 \leq z < 1.0$	44	2	6
$1.0 \leq z < 1.6$	15	5	1

^aThe F07 cluster list contains 72 clusters, 7 of which have $z < 0.20$ and are thus not included in this study.

^bIlbert et al. (2009)

^cWe use the area covered by the I09 photo-z catalogue, see Sect. 6.1.1.

the usable projected area (i.e. full cutout minus masked area) calculated. Each cluster for which galaxies cover less than 90% (50%) of the projected field is then rejected. The radius for the photo-z catalogue equals the radius to which the luminosity and richness measurements are made. The radius for the shear catalogue is the maximum radius for which shear measurements are made. The number of clusters rejected can be found in Table 7.1, and the corresponding details in Tables A.1 and A.2. Note that the clusters can have low coverage in both the photo-z and shear catalogues, i.e. the two rejection criteria are not mutually exclusive.

7.3 Measuring cluster properties: luminosity and richness

The optical luminosity and richness of a cluster are calculated using the redshift information and luminosity of individual galaxies. The photo-z catalogue by I09 provides 305 002 galaxies with $i < 25$, all of which are unmasked in all the optical bands and hence have good redshift estimates (see Sect. 6.1.5 for more details regarding this catalogue). It contains the aperture magnitude ($3''$ aperture, Sect. 4.3.1) with errors for each object in the u^* , B , V , g^+ , r^+ , i^+ , i' , z^+ , J , K bands. Note that there are two different i bands: Subaru i^+ is the main filter, but given that the i^+ -band image saturates at magnitude 22 for point sources, the CFHT i' is used when the object has no coverage with the Subaru filter.

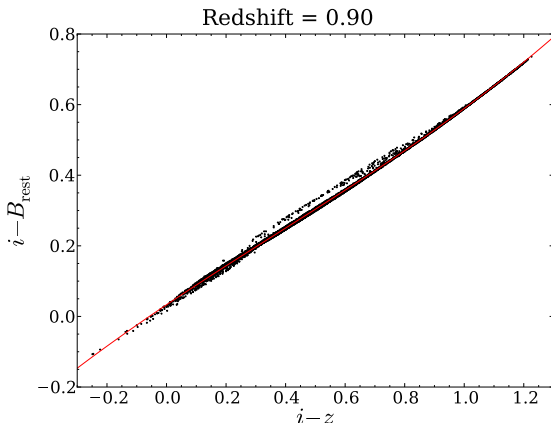


Figure 7.1: Typical colour-colour diagram used for calculating the apparent magnitude of an object, interpolated to the wavelength corresponding to the rest-frame B band (B_{rest}). The diagram is derived from galaxy SED templates used for calculating photometric redshifts, here shown at redshift 0.90. The red line shows that a 3rd order polynomial fits very well to the data, allowing the calculation of $f_{B_{\text{rest}}}$ from $(i - z)$ for each object at $z = 0.90$. The bimodality arises due to the galaxy SEDs used and corresponds to templates with emission lines. It affects only a small fraction of the galaxies on a negligible level (~ 0.02 mag). Similar plots are used for the whole redshift range (see text for details).

7.3.1 Calculating luminosities of individual objects

Calculating the luminosity of an object is done in several steps. Essentially, their magnitudes are transformed to a rest-frame system, taking redshifts, distances, and distance moduli into account, after which their luminosities can be obtained. We choose the B band as rest-frame filter, as we would like to keep as many galaxies as possible. The rest-frame V band might be a better choice, since its effective wavelength is further away from the 4000\AA break. However, we do not want to extrapolate outside our filter set, and given the filters provided in the catalogue, using V_{rest} would limit the sample to lower redshifts (see below).

Given that they are at different redshifts, we must first calculate each object's apparent magnitude interpolated to the wavelength corresponding to the rest-frame B band, $m_{B_{\text{rest}}}$. To obtain $m_{B_{\text{rest}}}$ the spectral energy distribution (SED) of the object must be taken into account. We do this by utilising galaxy SED templates (used for calculating photo-zs) to plot the relationship between the observed colour of the object, $(m_{f_1} - m_{f_2})_{\text{obs}}$, and the rest-frame colour, $(m_{f_1} - m_{B_{\text{rest}}})_{\text{fit}}$, for a given redshift (see Fig. 7.1). Here, $m_{f_1, \text{obs}}$ represents the object magnitude in the bluer filter (f_1) of the two and $m_{f_2, \text{obs}}$ the redder (f_2). Since we do not want to make one plot for each object (at each redshift), we utilise the constant $m_{f_1, \text{fit}}$ and include the individual object magnitudes in eq. (7.7) below. A fit to the colour-colour relation is then made to obtain $(m_{f_1} - m_{B_{\text{rest}}})_{\text{fit}}$, after which $m_{B_{\text{rest}}}$ is calculated according to

$$\begin{aligned} (m_{f_1} - m_{B_{\text{rest}}})_{\text{obs}} &= (m_{f_1} - m_{B_{\text{rest}}})_{\text{fit}} \\ m_{B_{\text{rest}}} &= m_{f_1, \text{obs}} - (m_{f_1} - m_{B_{\text{rest}}})_{\text{fit}}. \end{aligned} \quad (7.7)$$

Figure 7.1 shows an example plot for $i - z$ at redshift 0.90. Similar plots are derived for the whole redshift range in steps of $\Delta z = 0.01$. A 3rd-order polynomial fits nicely at all redshifts. The lowest redshift is set to $z_{\text{low}} = 0.15$, as the lowest cluster redshift is $z_{\text{cl}} = 0.20$ and the galaxies having $z < z_{\text{low}}$ are not affecting the richness/luminosity estimate of the lowest-redshift cluster (nor any of the higher-redshift ones). The upper redshift cut is $z_{\text{high}} = 1.65$, due to the filters and their effective wavelengths. The effective

Table 7.2: Filter choice with respect to redshift range for luminosity calculations. The redshift of an object decides which filters are used to calculate its luminosity.

Redshift	colour
0.15 – 0.57	$g^+ - r^+$
0.58 – 0.79	$r^+ - i^+$
0.80 – 1.17	$i^+ - z^+$
1.18 – 1.65	$z^+ - J$

Table 7.3: Number of galaxies in the photometric redshift catalogue.

Photo-z catalogue ^a	
Original	305 002
Only $0.15 \leq z < 1.65$	243 784
Cut wrt σ_z^b and $L_{B_{\text{rest}}}^*$	85 259
Cut wrt $L_{B_{\text{rest}}}^*$ only	111 667
Cut wrt σ_z only	163 083

^aIlbert et al. (2009)

^bAll errors in this table are given on the 95% confidence level

wavelength of the rest-frame B -band filter equals 4478\AA , increasing with redshift as $(1+z)$. To find the filters that match $(1+z)m_{B_{\text{rest}}}$ best we look at the efficient wavelengths of each filter. The filters that are closest in wavelength with respect to redshift are then used. Table 7.2 shows which filters were used in which redshift range.

This method works well up to the observed J band. However, between the J and the K filters there is a large gap of $\sim 9000\text{\AA}$, making the fits above the J -band wavelength rather unstable. We therefore only include galaxies in the redshift range for which we can obtain stable $m_{B_{\text{rest}}}$ estimates, meaning that we need to make an upper redshift limit at $z_{\text{high}} = 1.65$. The photo-z catalogue for which $0.20 \leq z < 1.65$ contains 243 784 galaxies (Table 7.3).

Since the objects are at different distances their apparent magnitudes $m_{B_{\text{rest}}}$, which is what we measure on the sky, must be transformed to their distance-independent absolute magnitudes $M_{B_{\text{rest}}}$,

$$m_{B_{\text{rest}}} - M_{B_{\text{rest}}} = 5 \log_{10} \left(\frac{D_{\text{lum}}}{10 \text{ pc}} \right) - 2.5 \log_{10}(1+z), \quad (7.8)$$

where D_{lum} is the luminosity distance (Sect. 1.7). The last term of eq. (7.8) corrects for the broadening of the rest-frame band as a function of redshift (part of the K correction, e.g. Hogg et al. 2002; see also Sect. 1.7).

From $M_{B_{\text{rest}}}$ we can finally calculate the rest-frame B -band luminosity

$$L_{B_{\text{rest}}} = 10^{-0.4(M_{B_{\text{rest}}} - M_{\odot, B_{\text{rest}}})} L_{\odot, B_{\text{rest}}}. \quad (7.9)$$

7.3.1a Filtering the photo-z catalogue

Since our photo-z catalogue is flux limited ($i < 25$), the average luminosity will increase as a function of distance. If not taken into account, this Malmquist bias will thus lead to an over-representation of luminous galaxies. We correct for this effect by selecting objects with luminosity

$$L_{B_{\text{rest}}} > 0.245(1+z)L_{B_{\text{rest}}}^* = 0.5(1+z) \times 10^{10} h_{70}^{-2} L_{\odot, B_{\text{rest}}} \quad (7.10)$$

(111 667 galaxies), where $(1+z)$ accounts for passive evolution, i.e. the evolution of luminosity as a function of redshift (see Sect. 7.5.1 for details). $L_{B_{\text{rest}}}^*$ is the characteristic galaxy luminosity where the power-law form of the Schechter luminosity function cuts off, here measured in the rest-frame B band (i.e. not changing with redshift). The Schechter luminosity function describes the number density of galaxies as a function of luminosity. The cut in eq. (7.10) represents the luminosity threshold out to which we can detect all existing galaxies. Application to the photo-z catalogue results in a sample complete over the whole redshift range.

In addition, to eliminate outliers from our photo-z catalogue we reject galaxies with $\sigma_z > 2 \times \text{median}(\sigma_z)$, where each object's redshift errors, σ_z , represent the 95% confidence level. The filtering is done in redshift bins according to cluster redshifts, such that $z_i \leq z_{\text{cl}} < z_{i+1}$, where each bin has a width of $\Delta z = z_{i+1} - z_i = 0.01$ and $\text{median}(\sigma_z)$ is calculated from the redshift errors of the clusters within each individual bin. The final photo-z catalogue used for calculating cluster richnesses and luminosities now contains 85 259 galaxies (see Table 7.3). To make sure this redshift cut does not introduce a bias, we repeated the analysis including the galaxies initially rejected in this step. The results have higher errors due to the inclusion of noisy data, but they all agree within error bars with those excluding the high redshift error galaxies. The results from the $M - L_{B_{\text{rest}}}$ fit for both the filtered and the non-filtered catalogues agree well when using the optically selected cluster list. For the X-ray clusters the results still coincide within error bars, albeit with larger scatter. The slopes of the $M/L_{B_{\text{rest}}}$ vs. redshift fits agree well within error bars when applied to both the filtered and the non-filtered catalogues. The same holds for the relation between X-ray temperature and optical luminosity ($T_X \propto L_{B_{\text{rest}}}$). For the remainder of this chapter we will drop the “rest” subscript notation, such that all references to “ B ” refer to the rest-frame B band.

7.3.2 Richness

The richness of a galaxy cluster is defined as the total number of cluster galaxies within a given radius and with luminosities above a certain threshold. From simulations the richness is also described as the halo occupation number, HON, the mean number of galaxies within haloes of a given mass. To differentiate between cluster and field galaxies, extensive spectroscopic measurements are required. However, given that we have no spectroscopic information for our clusters, we can use statistical background subtraction to determine the overdensity of cluster galaxies instead. This method also provides a consistent technique for each cluster, avoiding biases between clusters that might arise from a poor determination of the selection function (e.g. Muzzin et al. 2007).

Firstly, we must define the redshift slice around the cluster redshift, z_{cl} , that contains the cluster members. We choose $z_{\text{cl}} \pm 0.025(1+z_{\text{cl}})$, where all galaxies whose redshift

estimate lies inside the redshift slice are chosen. In addition, all galaxies whose redshift errors cover the cluster redshift are also chosen as possible cluster members, i.e. those where $z - \sigma_z^- \leq z_{\text{cl}} \leq z + \sigma_z^+$. The redshift errors represent the upper and lower 95% confidence level. Note that we initially calculated the cluster richness using early-type galaxies (as defined by I09) only. However, given that above $z \sim 1.0$ almost no such galaxies could be detected, we resumed using all galaxies.

Secondly, the above cluster member selection must be done within a given physical radius that approximately matches the cluster size. Since the clusters in both lists all have masses within a small range this will not introduce a bias. We choose $r_{\text{cl}} = 1.0$ Mpc after testing different radii (see Sect. 7.3.2a). All galaxies within this radius fulfilling the above redshift criteria are cluster member candidates.

Finally, since we have no spectroscopic information it is impossible to discern between cluster members and field galaxies in the process described above. We must therefore perform a field galaxy subtraction from the resulting number count. Assuming field galaxies are randomly spread out over the entire field, we can apply the same cluster member selection method as above to random field positions (at $z = 0.2$ we can place 64 such apertures in the field whereas at $z = 1.60$ the field fits 440 apertures) and subtract the mean value from the initial cluster member count to obtain the field-galaxy subtracted cluster richness. Looking at the distribution of number of galaxies in each aperture we see that it is close to Gaussian with a fairly low high-luminosity tail, justifying this assumption. The error of the richness can be estimated by the standard deviation of the field-galaxy richnesses.

7.3.2a Testing the cluster radius

The smaller the cluster radius, the more reliable the field galaxy correction becomes, i.e. more sensitive to the cluster galaxies. With increasing aperture radius the absolute fluctuation of the number of galaxies inside the aperture will become larger than the number of cluster members, thus the correction will be less reliable. To investigate this effect we therefore calculate the richness of all the optical clusters within $r_{\text{cl}} = 0.3$ Mpc in addition, and plot them as a function of the richness within $r_{\text{cl}} = 1.0$ Mpc.

The choice of cluster radius, r , is a trade-off between our aims to reduce the sensitivity to cluster substructure (leading to larger r), and reduce the noise from the field subtraction (leading to smaller r). Doing this test we find that $r_{\text{cl}} = 1$ Mpc provides good results. Similarly, a fixed cluster radius has been employed in an earlier study by Hoekstra et al. (2002) for clusters in the redshift range $0.2 \leq z \leq 0.8$. In Fig. 7.2 we additionally compare the richness estimates within $r_{\text{cl}} = 1$ Mpc and $r_{\text{cl}} = 0.3$ Mpc. Here we find a good correlation with $N_{\text{gal}}(< r_{\text{cl}} = 1 \text{ Mpc}) \sim 2.5 \times N_{\text{gal}}(< r_{\text{cl}} = 0.3 \text{ Mpc})$. The relative richness errors are only weakly increased for $r_{\text{cl}} = 1$ Mpc, while it provides a better match to the typical cluster virial radius, confirming our choice.

7.3.3 Luminosity

The total luminosity of a cluster is measured using the same procedure as for calculating the cluster richness. That is, the individual luminosities of each galaxy within an aperture are added. Then the same routine is applied to random fields, and the mean of these aperture field-galaxy luminosities is subtracted from the original cluster luminosity.

7.3. Measuring cluster properties: luminosity and richness

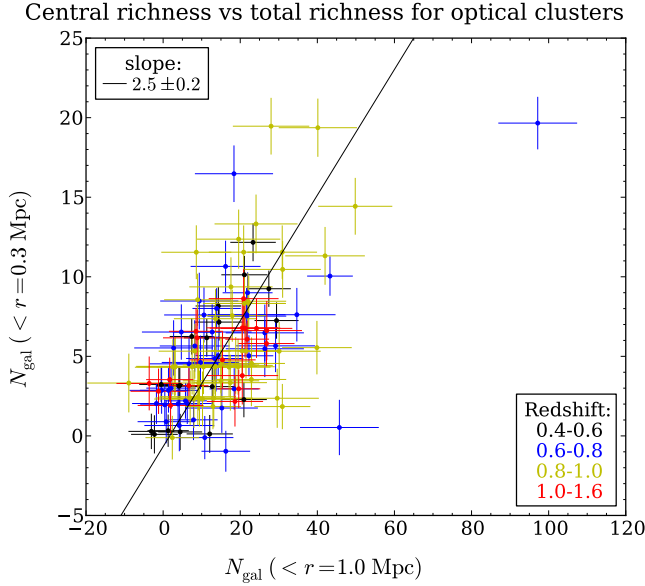


Figure 7.2: Central richness, calculated within a cluster radius of 0.3 Mpc, vs. total richness, calculated within $r_{cl} = 1$ Mpc, for optical clusters. Colours indicate cluster redshifts. The slope indicates the fit between the total richness and the central richness, $N_{total} \sim \text{slope} \times N_{central}$. Note that the seemingly negative luminosities arise due to the background subtraction of faint clusters, see text for details.

The error of the cluster luminosity is more complex than for the richness calculations. Using the individual galaxy luminosity errors as basis, we calculate the error of the total cluster luminosity as

$$\sigma_{L,cl} = \sqrt{\sigma_{L,cl,orig}^2 + \frac{\langle \sigma_{L,fg,orig}^2 \rangle}{n} + \text{var}(L_{fg})}. \quad (7.11)$$

The first term refers to the error of the unsubtracted (original) total cluster luminosity. If σ_{LB} is the error of each individual galaxy contributing to this original luminosity count, then $\sigma_{L,cl,orig}^2 = \sum \sigma_{LB}^2$ (see below for a deduction of σ_{LB}). The second term handles the same errors as the first term, but for the field-galaxy luminosity subtraction. Since the mean of the aperture field-galaxy luminosities is used when calculating the cluster luminosity, the variance of this mean must be included in its error calculations. Here, n equals the number of field apertures used in the mean, and $\langle \sigma_{L,fg,orig}^2 \rangle$ is the mean of each aperture field-galaxy luminosity error estimate. In addition to these two terms, one must also take into account the variance between the total luminosities calculated for field galaxies in each aperture.

Each individual galaxy's luminosity error is calculated from its magnitude error, given in the photo-z catalogue. We first define

$$x(m_{f_1}, m_{f_2}) = m_{f_1} - m_{f_2} \quad (7.12)$$

$$g(x) = \sum_{i=0}^3 a_i x^i, \quad (7.13)$$

where $m_{f_{1,2}}$ denote the apparent magnitudes in the bluer (f_1) and redder (f_2) filters, and $g(x)$ is the 3rd order fit to $(m_{f_1} - m_B)_{fit}$, see Sect. 7.3.1 for details. From eq. (7.7) we

then have

$$m_B(m_{f_1}, m_{f_2}) = m_{f_1} - g[x(m_{f_1}, m_{f_2})] \quad (7.14)$$

$$\begin{aligned} \sigma_{m_B}^2 &= \left(\frac{\partial m_B}{\partial m_{f_1}} \sigma_{m_{f_1}} \right)^2 + \left(\frac{\partial m_B}{\partial m_{f_2}} \sigma_{m_{f_2}} \right)^2 \\ &= \left(1 - \frac{\partial g}{\partial x} \frac{\partial x}{\partial m_{f_1}} \right)^2 \sigma_{m_{f_1}}^2 + \left(-\frac{\partial g}{\partial x} \frac{\partial x}{\partial m_{f_2}} \right)^2 \sigma_{m_{f_2}}^2 \\ &= \left(1 - \frac{\partial g}{\partial x} \right)^2 \sigma_{m_{f_1}}^2 + \left(\frac{\partial g}{\partial x} \right)^2 \sigma_{m_{f_2}}^2 . \end{aligned} \quad (7.15)$$

Equation (7.8) yields

$$\sigma_{M_B}^2 = \left(\frac{\partial M_B}{\partial m_B} \sigma_{m_B} \right)^2 = \sigma_{m_B}^2 , \quad (7.16)$$

where we ignore the uncertainty of the redshift estimate. We can do this because the redshift uncertainties are dominated by the magnitude errors $\sigma_{m_{f_{1,2}}}$ which are already incorporated in the photo- z estimation procedure. From eq. (7.9) we have

$$\begin{aligned} \sigma_{L_B}^2 &= \left(\frac{\partial L_B}{\partial M_B} \sigma_{M_B} \right)^2 \\ &= \left(L_B \ln 10 \cdot (-0.4) \right)^2 \sigma_{M_B}^2 . \end{aligned}$$

The final expression is now given by

$$\sigma_{L_B} = 0.4 L_B \ln 10 \sqrt{\left(1 - \frac{\partial g}{\partial x} \right)^2 \sigma_{m_{f_1}}^2 + \left(\frac{\partial g}{\partial x} \right)^2 \sigma_{m_{f_2}}^2} . \quad (7.17)$$

Figure 7.3 shows the cluster luminosities as a function of redshift for both cluster lists. The luminosities have been corrected for passive evolution, see Sect. 7.5.1 for more details.

7.4 Measuring the cluster mass

To estimate the cluster masses we use weak gravitational lensing (see Chap. 2 for details). For this purpose we use the shear catalogue by Schrabback et al. (2010), which is described in greater detail in Sect. 6.1.3. The shear catalogue is matched with the I09 photo- z catalogue assigning redshifts to 194 440 galaxies. 256 944 galaxies remain without redshift information. Section 6.4.1 describes in detail how the redshift distribution of galaxies without photo- z s is modelled.

The Lyman-break technique provides additional, more accurate redshift information regarding high-redshift galaxies, see Sect. 6.1.4 for details. We therefore apply Lyman-break galaxy (LBG) information when possible, provided by the Hildebrandt et al. (2009) LBG catalogue. As in Chap. 6, we discard the r -band dropouts and use only the u - and g -dropouts (see Sect. 6.3 for a justification of this decision). These dropout catalogues are then matched with the shear catalogue, yielding a total of 10 177 galaxies in the LBG

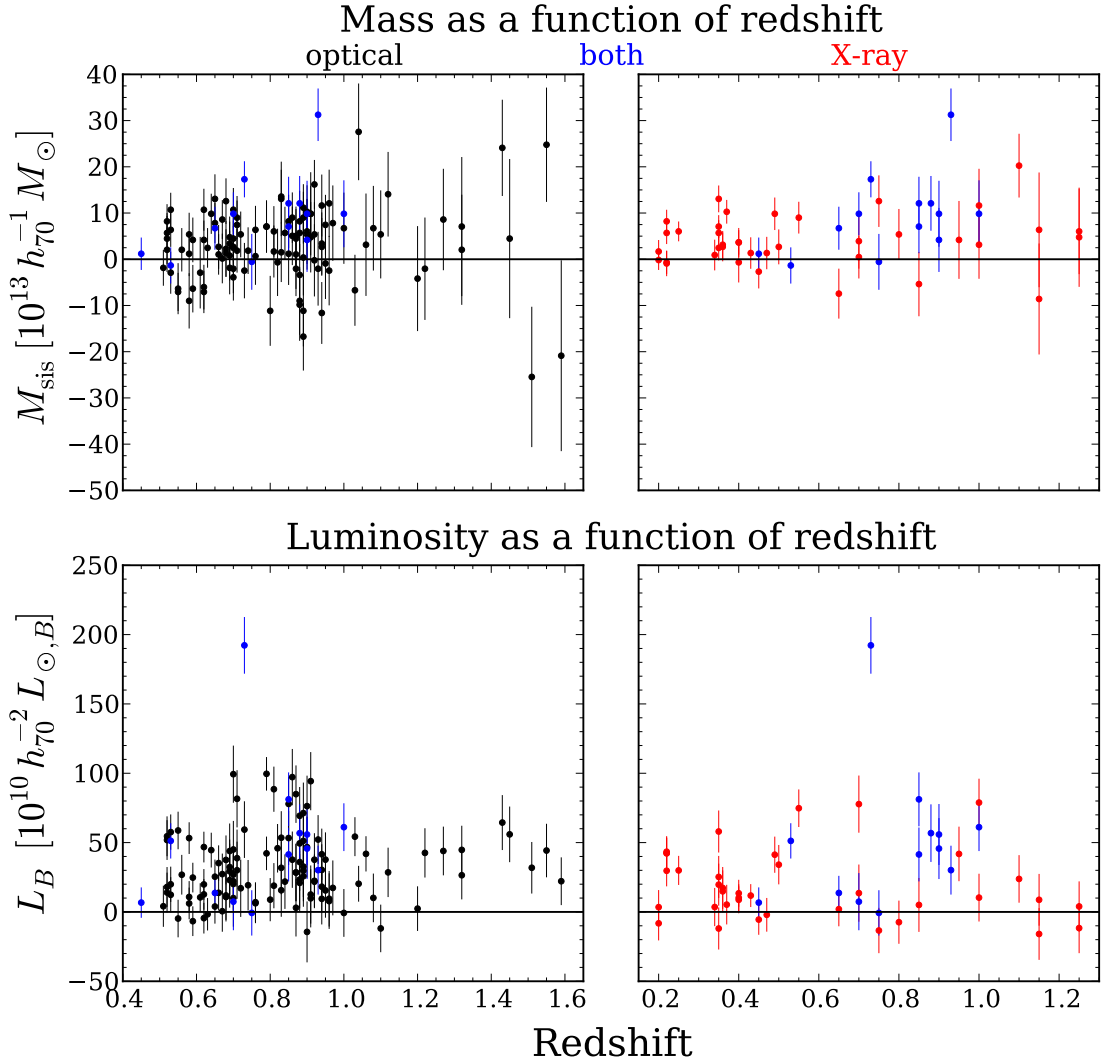


Figure 7.3: Mass (*top*) and luminosity (*bottom*) as a function of redshift for the optical (*left*) and X-ray (*right*) cluster lists. The luminosities have been corrected for passive evolution. Clusters included in both cluster lists are plotted in blue. Note that the redshift range is different between the two samples. Also, the apparent rise in average luminosity towards higher redshifts for the optically selected clusters is not a selection effect since our photo- z sample is complete (see Sect. 7.3.1a).

shear catalogue (Table 6.1). The photo- z and nonphoto- z shear catalogues are mutually exclusive, however they both contain LBGs. Sect. 6.1.4 gives a detailed description of the Lyman-break technique and galaxies, along with an overview of the LBG catalogue.

To increase the relative lensing efficiency (Sect. 2.2.4) we follow the weighting approach described in Sect. 6.2, and scale the shear of the source galaxies to $z = 3.0$. Given that we do not stack the clusters but look at trends between individual clusters we do not scale the clusters to a common cluster redshift. To avoid cluster galaxy contamination we also remove sources (from the photo- z shear catalogue) that have $z_s < z_{\text{cl}} + 0.2$ (Abate et al. 2009).

The shear profile of each cluster is modelled by calculating the tangential shear (Sect. 2.3) for all background galaxies within $0.1 \text{ Mpc} < r < 2.0 \text{ Mpc}$, see Sect. 6.2.1 for a discussion regarding the selection of $r_{\text{min}} = 0.1 \text{ Mpc}$. We then fit the singular isothermal sphere (SIS) model to obtain a mass estimate of the cluster. The SIS model describes an isothermal, spherical, and axially symmetric mass distribution (Sect. 2.4.2). The more complex Navarro, Frenk, & White (1997, 1995) profile is normally considered more applicable in describing the mass distribution of galaxy clusters (Sect. 2.4.3). However, since the clusters in this study span a wide redshift range ($0.2 \leq z \leq 1.6$), they are likely to comprise very different radial profiles. The signal-to-noise (S/N) for each cluster is in addition very poor, and we therefore choose to fit the simplest profile.

The SIS model is characterised by the Einstein angle, θ_E , which is the only free parameter once the cluster centre is fixed. It can be written as a function of the velocity dispersion, σ_v , see eq. (2.79). Using eq. (2.80) and assuming the underlying potential of the lens is well approximated by an SIS profile, a direct measurement of θ_E (and hence also σ_v) of the lens comes from measuring the shear inside a projected radius θ . An estimate of the mass of the lens, M_{sis} , within a three-dimensional radius r can be obtained from inserting σ_v into

$$M_{\text{sis}}(r) = \frac{2\sigma_v^2 r}{G}. \quad (7.18)$$

Note that the best-fit θ_E is found using the minimum χ^2 method, which also provides upper and lower error bars for the fit.

The upper panel of Fig. 7.3 shows the cluster masses as a function of cluster redshifts. Negative masses arise for less massive clusters, where the measured lensing signal is dominated by noise. For the optical clusters (left plot), problems with cluster centring is also likely to dilute the lensing signal. Finding the cluster centre is a difficult task (Sect. 5.5.3). The three-dimensional reconstruction method used by Z08 does not select a single galaxy as cluster centre, instead the luminosity weighted peak of the galaxy distribution is chosen. Although this method might work very well for massive clusters, with less massive clusters it might cause problems, in which case choosing the brightest cluster galaxy (BCG) as cluster centre is more robust.

One example that picking the wrong cluster centre can be catastrophic comes when two clusters are spatially lying very close. They are likely to be mistaken as being only one cluster, and hence will produce a cluster ‘‘centre’’ lying in the middle between the two clusters. Both clusters will contribute to the luminosity measurements, whereas the lensing signal will be diminished. Cluster number 92 is one such example, where it is clear from looking at the FITS images that it does in fact consist of two clusters, each lying on opposite sides of the nominal cluster centre. Before this was revealed, cluster number 92 showed up with a luminosity of $\sim (160 \pm 20) \times 10^{10} h_{70}^{-1} L_{\odot,B}$ but with zero mass.

7.5 Results and discussion: mass-to-light ratio and mass-richness relation

Note that fits in this chapter are made using orthogonal distance regression¹, which takes the errors in both input and response variables into account. In case the input variable does not have errors associated with it (e.g. the M/L_B vs. redshift fit in Sect. 7.5.3), the fit is made using the least squares method² taking only the errors in the response variable into account.

7.5.1 Correction for passive evolution

The mass-to-light ratio may depend on mass (see below) and redshift. The redshift dependence is due to passive evolution in the cluster galaxy luminosities, and scales with cluster luminosity as $\propto(1+z)$. Passive evolution describes how luminosities and colours of early-type galaxies and stellar populations evolve in the absence of star formation. The simplest form of passive evolution describes the evolution of a single age stellar population formed at a given redshift, but can also allow for complex star formation histories of individual galaxies as well as morphological evolution in the galaxy population (e.g. van Dokkum & Franx 2001).

Without submerging into a detailed study of the spectral evolution of stellar populations and star formation in (cluster-) galaxies, one can correct M/L for passive evolution by simply dividing the cluster luminosities by $(1+z)$. The evolution of the brightest galaxies is restricted to a minimal passive evolution, as has been established by ground-breaking work on faint galaxy evolution (e.g. Koo & Kron 1992; Ellis et al. 1996). The simple correction above is hence also justified by the cut with respect to L^* applied to the photo- z catalogue in Sect. 7.3. By choosing only bright galaxies, where luminosity and colour evolution play a smaller role than they do for less luminous galaxies, complications due to differential evolution are substantially minimised (e.g. Carlberg et al. 1999). For the remainder of this chapter all optical luminosities are corrected for passive evolution unless otherwise stated.

7.5.2 Luminosity vs. richness

First, we look at the connection between the luminosities and the richnesses of the clusters. Given that both measurements for the same cluster are made from the same galaxies, we expect a tight relation between the two. Systematic errors in the cluster richness and luminosity calculations can be revealed by such a comparison. Figure 7.4 shows the luminosity-richness relation for each of the cluster lists. The expected strong relation is apparent, independent of redshift.

¹Python's `scipy.odr` package: <http://docs.scipy.org/doc/scipy/reference/odr.html>

²using Python's `scipy.optimize.leastsq` package:
<http://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.leastsq.html>

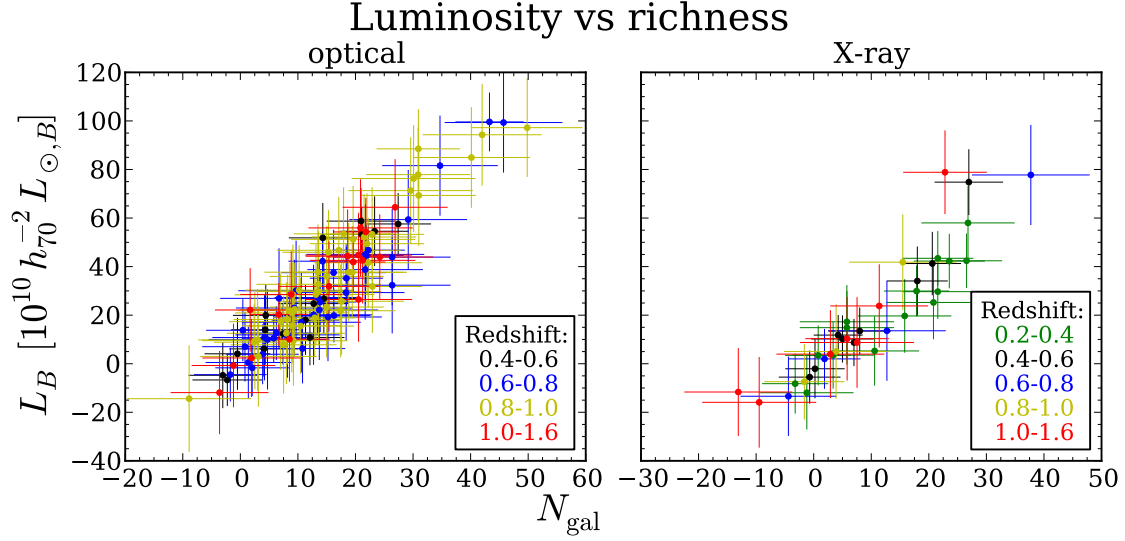


Figure 7.4: Luminosity-richness relation for optical (*left*) and X-ray (*right*) clusters. Colours indicate cluster redshifts.

7.5.3 Mass-to-light ratio

To constrain the mean mass-to-light ratio, regardless of redshift, we fit the mass as a function of luminosity for all clusters in the same cluster list as

$$M = \langle M/L_B \rangle L_B . \quad (7.19)$$

Figure 7.5 shows M/L_B for both lists, where the red line denotes the fit to X-ray clusters and the black line the fit to the optical clusters. Note that the clusters plotted in blue are clusters included in both cluster lists and are thus included in both fits. Our results yield

$$M/L_B = (115 \pm 13) h_{70} M_{\odot}/L_{\odot,B} \quad (\text{optical}) \quad (7.20)$$

$$M/L_B = (146 \pm 21) h_{70} M_{\odot}/L_{\odot,B} \quad (\text{X-ray}) . \quad (7.21)$$

To quantify its significance on the fit, we also investigate the influence from the single, very luminous cluster in the upper right part of the figure (#81/#106 optical/xray). Fitting the same relation as above to the samples while excluding this cluster yields $M/L_B = (125 \pm 15) h_{70} M_{\odot}/L_{\odot,B}$ (optical) and $M/L_B = (183 \pm 28) h_{70} M_{\odot}/L_{\odot,B}$ (X-ray). These results both agree with the corresponding original mass estimates within error bars. We see that the outlier cluster affects the fit to the X-ray clusters more than that of the optically selected clusters, which is also expected due to the smaller number of clusters in the X-ray list.

Next we want to break down the above fit into redshift bins and investigate the M/L_B redshift dependence. Figure 7.6 shows the same as Fig. 7.5, but now the clusters have been divided into redshift bins before fitting. The redshift-independent fit is indicated with dashed lines in all bins. Finally, Fig. 7.7 shows the best-fit M/L_B in each redshift bin as a function of the mean of the cluster redshifts in the corresponding bin. The M/L_B error bars indicate the errors of the M vs. L_B fit. The width of each redshift bin is indicated in

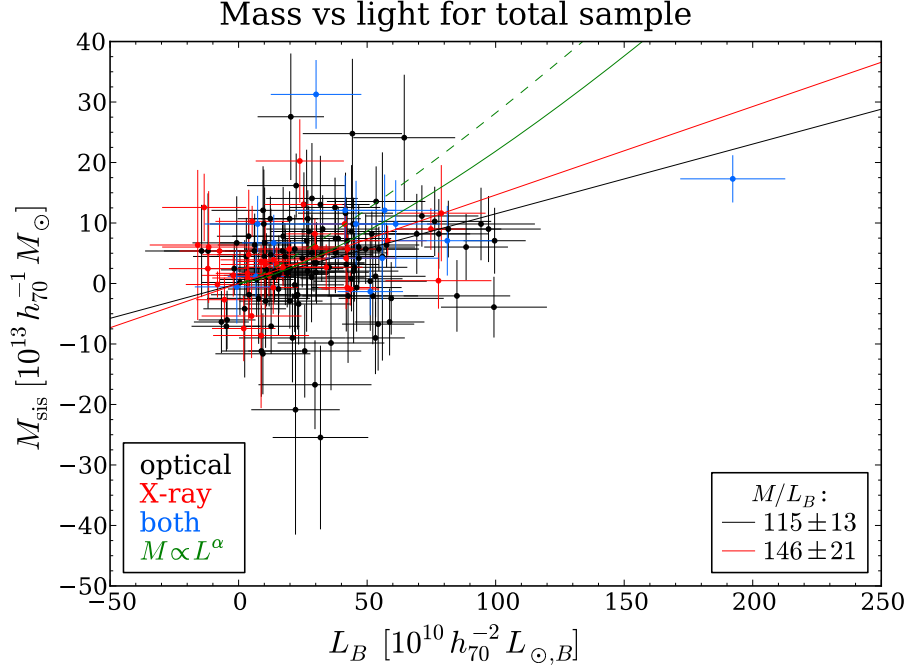


Figure 7.5: Mass as a function of optical luminosity for X-ray (red and blue) and optical (black and blue) clusters. The mass-to-light ratio for the two cluster samples are shown as solid lines (red: X-ray, black: optical). The green lines indicate power-law relations between M and L_B obtained by Girardi et al. (2002) (solid line; eq. (7.22)) and Muzzin et al. (2007) (dashed line; eq. (7.25)). Note that the relation obtained by Muzzin et al. (2007) is fit using K -band data.

grey. The M/L_B error bars (but not the redshift binwidths) are taken into account when fitting M/L_B vs. redshift. The number of clusters in each redshift bin can be found in Table 7.1.

There are several approaches to fitting M/L_B vs. redshift. Binning the data is usually not the best solution due to the suppression of information. However, in this case fitting individual M/L_B values as a function of redshift will introduce a bias, originating from applying a non-linear operation (dividing) to a noisy quantity (L_B) which causes the expectation value to change. Binning in redshift is therefore necessary to avoid this bias. There are several ways of binning the data, yet we find that the only unbiased method that correctly takes both M and L_B errors into account while also giving more massive and/or luminous clusters higher weights is to fit M vs. L_B in redshift bins as described above.

Figure 7.7 shows two plots in which M/L_B has been fit as a function of redshift. Both figures comprise the total data sample, in which the clusters have been divided in redshift bins of width $z = 0.2$ before fitting M vs. L_B . The only difference between the two figures is the starting point – the top plot starts at $z = 0.2$ and the bottom at $z = 0.1$ – hence the bins have been shifted by $z = 0.1$ between the two fits. In this way we test the impact of different binning schemes with respect to the results. For the X-ray selected cluster sample there is no difference between the two fits. The slopes are consistent with zero, and there is no significant indication for an evolution of M/L_B with redshift. Hoekstra

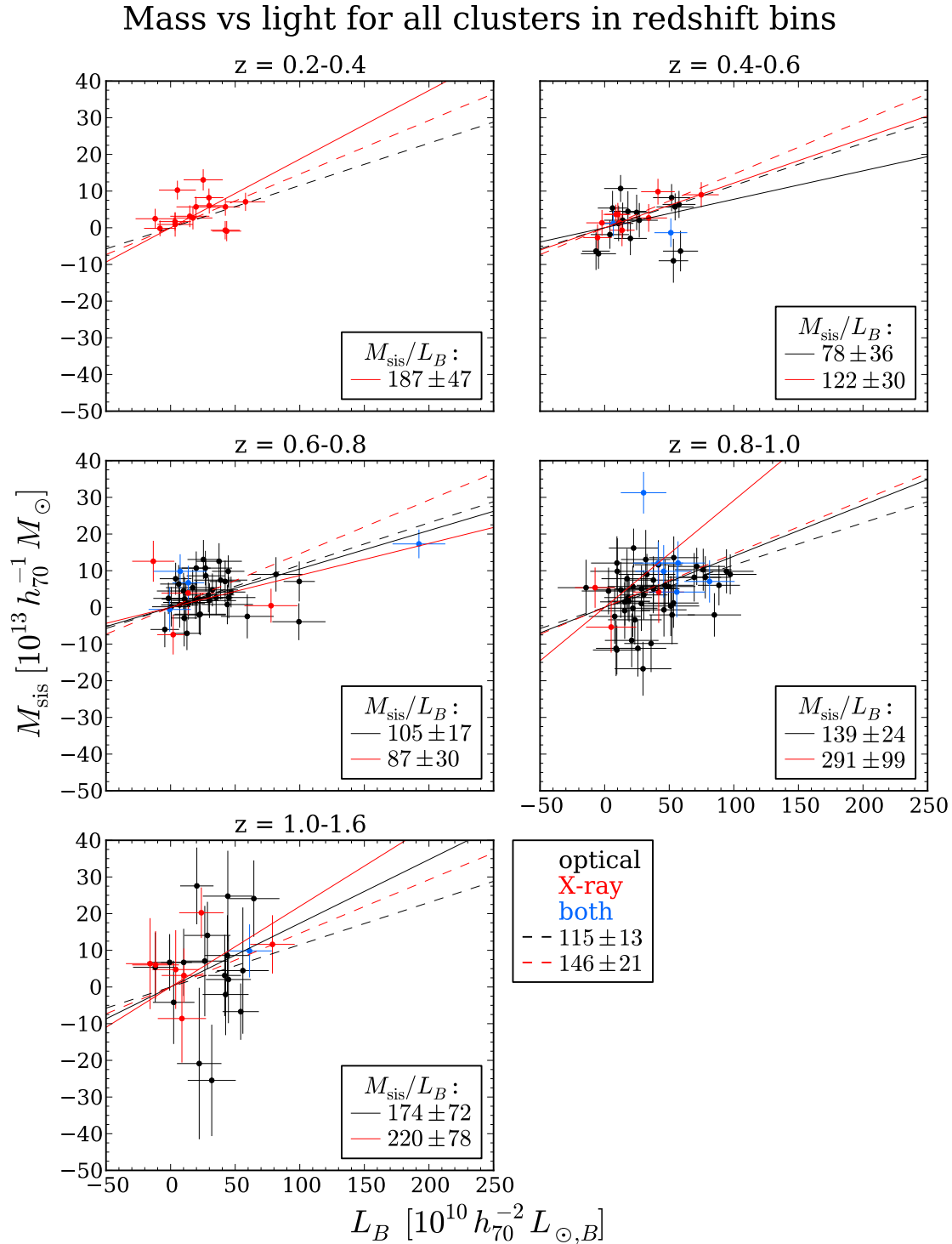


Figure 7.6: Mass vs. optical luminosity for X-ray (red and blue) and optical (black and blue) clusters. The clusters have been divided into redshift bins and the mass-to-light ratio fitted in each bin (solid lines). The M/L_B ratios for the total datasets are shown as dashed lines, and their numbers given in the figure legend. The clusters presented in blue are common objects and hence included when fitting M/L_B to both the X-ray clusters as well as to the optically selected clusters.

7.5. Results and discussion: mass-to-light ratio and mass-richness relation

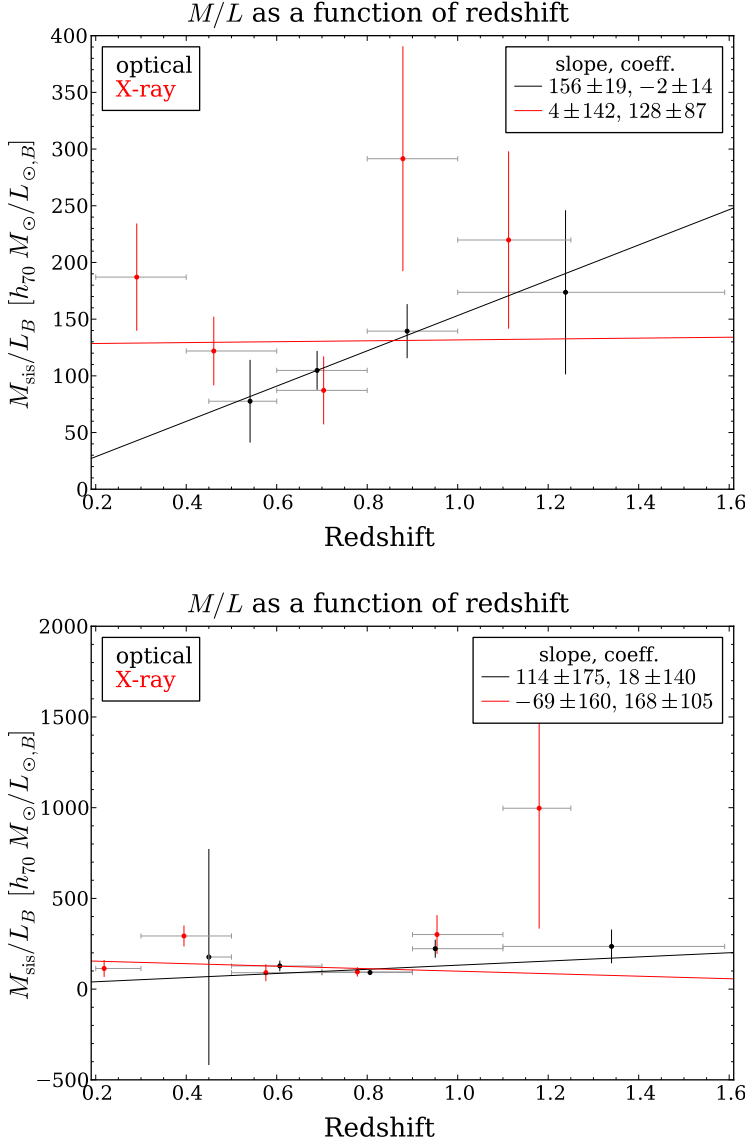


Figure 7.7: M/L_B as a function of redshift for all clusters. The M/L_B ratios have been fitted in redshift bins, and the error bars indicate the errors of the fit. The mean redshift in each bin is plotted, and the corresponding binwidth is indicated in grey. The M/L_B errors are taken into account when fitting M/L_B vs. redshift. The best-fit values are given in the legends. The data points in the *top* plot have been binned as throughout this thesis. We also include the *bottom* plot where the binning has been shifted by $z = 0.1$ to show that the significance of the fit to the optically selected clusters depends on the binning, see the text for more details. Both figures contain the same data, binned with binsize $z = 0.2$. Note that the scale on the y -axis is different in the two plots.

et al. (2002) reach a similar conclusion in their weak lensing study of 4 X-ray luminous clusters within $0.2 \leq z \leq 0.8$, in which the M/L_B evolution as a function of redshift is consistent with the corresponding luminosity evolution.

For the optically selected clusters the different binning matters more. The slopes are similar, though the significance of each fit is very different between the two figures. Overall this cluster sample shows a weak indication for an increasing M/L_B with redshift, but at very low significance given the uncertainties. This apparent trend has one of three likely causes; either the evolution is real, it originates from selection effects, or there are spurious detections in the cluster catalogue. The selection effects arise from the fact that we are dealing with a volume-limited survey, in which the probability of finding massive clusters varies as a function of redshift. Due to having a small co-moving volume at low redshifts, we do not expect to find many (if any at all) massive low-redshift clusters in the

COSMOS field. Moving towards higher redshifts this expectation value increases, peaking at $z \sim 0.5 - 1.0$. However, from structure growth we know that clusters start forming at $z \sim 1.0$, increasing in mass towards lower redshifts. Combining these two factors, we expect the most massive clusters to appear at $z \sim 0.5 - 0.8$. Looking at the top left panel of Fig. 7.3 we see that the most massive clusters in the optically selected cluster list do indeed appear at $z \gtrsim 0.6$. Given that the optical cluster list starts at $z \sim 0.5$, the selection effects due to volume should not impact our results regarding the evolution of M/L_B with respect to redshift. What might affect the results more, however, comes from spurious cluster detections in the cluster catalogue. These effects are larger at lower redshifts where lower-mass systems dominate, and will artificially bring down the M/L_B ratio with decreasing redshift. Finding the “correct” cluster centre (see Sect. 7.4) is also a concern in this matter, in that an off-centring will lead to a dilution of the lensing signal, lowering the lensing mass. As before, this effect also has a higher impact at lower redshifts. Given that the X-ray sample shows no evolution with redshift, the effect from spurious detections and cluster centring problems are the most likely explanations.

Muzzin et al. (2007) study the luminosity-mass relation for clusters from $z \sim 0.3$ to $z = 0$ and find little change. This was also found by Bahcall & Comerford (2002), who got consistent results (within 1σ) between clusters at $z < 0.1$ and $0.15 \lesssim z < 0.3$. A comparison between studies of very low-redshift clusters and our results from the lowest-redshift bin ($0.2 \leq z < 0.4$) is therefore justified.

Using semi-analytical modelling, Kauffmann et al. (1999) predict that M/L increases with mass with roughly the same slope in both B and I bands. The slope of the $M/L - M$ relation is also found observationally to be the same in all B, V, R, K bands (see the introduction of Popesso et al. 2007; and references therein). In addition, Popesso et al. (2007) find that the slopes of the best-fit $M - L$ and $M/L - M$ relations do not depend on which of the SDSS g, r, i, z bands are used. A comparison between our results and those given in other passbands is therefore valid.

Parker et al. (2005) investigate the mass-to-light ratio of galaxy groups. They use weak gravitational lensing to measure the tangential shear profile of the groups, to which an isothermal sphere is then fitted. Since groups are smaller than clusters, the weak lensing signal can only be measured around a stacked version of the galaxy groups (similar to what is done in Chap. 6). The mass is calculated within a radius of $1.43 h_{70}^{-1}$ Mpc. The average number of group members is 3.9 galaxies and the median redshift for all the groups is $z_{\text{med}} = 0.323$. The luminosities of the sample have not been corrected for evolution. For this group sample Parker et al. (2005) obtain $M/L_B = (130 \pm 20) h_{70} M_{\odot}/L_{\odot,B}$. In order to compare our results to those of Parker et al. (2005), Fig. 7.8 shows M_{sis} as a function of L_B in redshift bins in which the cluster luminosities have not been corrected for passive evolution. We see that using clusters in the redshift range $0.2 \leq z < 0.4$, i.e. close to their z_{med} , our results of $M/L_B = (146 \pm 36) h_{70} M_{\odot}/L_{\odot,B}$ are in good agreement with those of Parker et al. (2005).

Parker et al. (2005) also split their sample into two subsamples; one containing poor groups ($\sigma_v < 190 \text{ km s}^{-1}$) and one containing rich groups ($\sigma_v > 190 \text{ km s}^{-1}$). The latter is of interest here, as we are considering mostly clusters towards the low-mass end, the smallest clusters resembling groups more than actual clusters. The mean N_{gal} for their rich group sample is 4.2, with a median redshift of $z_{\text{med}} = 0.360$. Their mass-to-light results for this sample yields $M/L_B = (195 \pm 29) h_{70} M_{\odot}/L_{\odot,B}$, still consistent with our

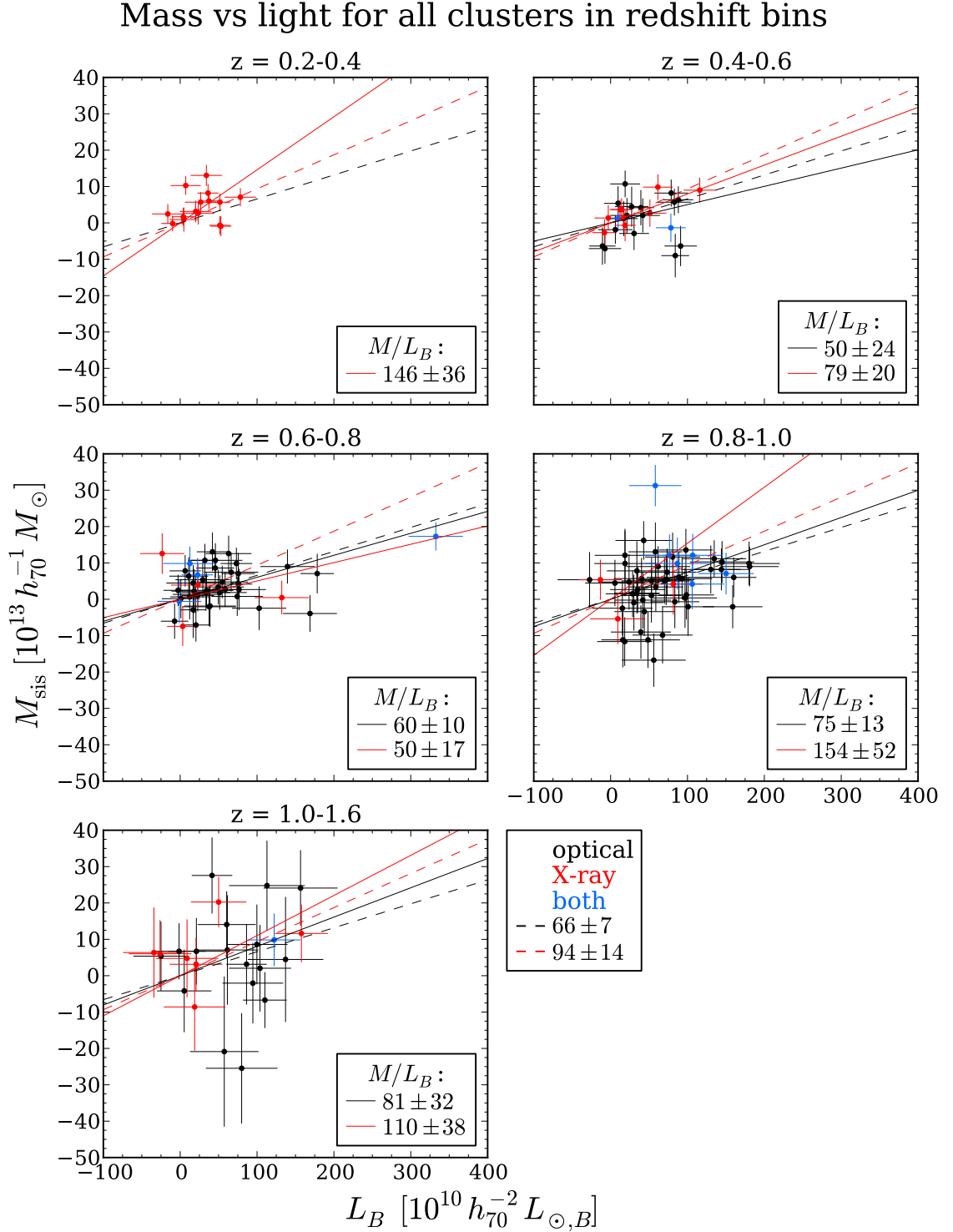


Figure 7.8: Mass as a function of uncorrected (optical) luminosity for X-ray (red and blue) and optical (black and blue) clusters. In order to compare our results to those of Parker et al. (2005), the luminosities in this figure have not been corrected for passive evolution, and differ in this respect to Fig. 7.6. See the Fig. 7.6 figure text for more details.

$0.2 \leq z < 0.4$ results.

Girardi et al. (2002) explore a sample of galaxy groups and clusters, comprising systems from within a large span in richness, all at low redshifts ($z_{\text{cl}} \leq 0.15$). They investigate the mass-luminosity relation by fitting a power law, and find that

$$\frac{M}{M_{\odot}} = 10^{-1.596 \pm 0.381} \left(\frac{L_B}{L_{\odot, B_j}} \right)^{1.338 \pm 0.033} (0.7 h_{70})^{-1.676 \pm 0.066} . \quad (7.22)$$

While fitting the same relation to our data it became clear that the data points are too spread and the uncertainties too large to constrain this type of fit. Since the fit was so unstable we did not pursue this approach further. However, when plotting eq. (7.22) on top of our data points in Fig. 7.5 (green solid line), we see that this relation fits well with our data.

Popesso et al. (2007) look at 217 galaxy clusters in the Sloan Digital Sky Survey with $\langle z \rangle = 0.1$, investigating the relation between M_{200}/L_r and mass (where $M_{200} \equiv M(r_{200})$), and finding $M_{200}/L_r \propto M_{200}^{0.18 \pm 0.04}$. Since individual M/L_B are used for fitting $M/L_B \propto M^{\alpha}$, we could not apply this method to our data without biasing our results (as explained earlier in this section, page 117). This bias can be overcome by fitting M vs. L_B in mass bins, however as stated above, we can not rule out a correlation between M/L and redshift. Hence ignoring a possible redshift dependence will also bias our results. The ideal solution would therefore be to first divide the data in redshift bins, and then fit M as a function of L_B in mass bins. Nevertheless, to yield viable results, this approach requires a much larger sample than is available at this point. With 217 clusters, Popesso et al. (2007) have a statistically much more robust sample compared to our lowest- z bin containing only 13 clusters. In addition, many of our clusters have low masses and resemble groups more than clusters. Trying to model a noisy, low-mass cluster/group sample with lensing naturally leads to a fraction of negative masses, to which the power-law cannot be fit. Fitting this function to low-mass clusters does not make sense either, as we do not expect $M/L = 0$ at $M = 0$ for $\alpha > 0$ (which the power-law assumes).

Bahcall & Comerford (2002) study galaxy groups and clusters ranging from poor to rich systems in the redshift range $0.02 \lesssim z \leq 0.8$, investigating how M/L_V depends on cluster overdensity, represented by the observed temperature (T_X) of the cluster. To calculate M/L_V they use mass estimates from the literature, which have been determined using both gravitational lensing and X-ray measurement methods. They find that M/L_V increases as a function of T_X as

$$M/L_V = (121 \pm 20) \left(\frac{k_B T_X}{1 \text{ keV}} \right)^{0.30 \pm 0.08} h_{70} M_{\odot} / L_{\odot, V} . \quad (7.23)$$

Due to the same reasons as for fitting $M/L_B \propto M^{\alpha}$ (see above), we could not fit M/L_B as a function of T_X to compare to the above results. We could, however, calculate the average temperature of our cluster sample and insert it into eq. (7.23) to compare to our M/L_B results. The mean temperature of our X-ray clusters within $0.2 \leq z \leq 0.8$ is $\langle k_B T_X \rangle = 1.17 \text{ keV}$ (for the total sample $\langle k_B T_X \rangle = 1.28 \text{ keV}$), yielding

$$\begin{aligned} M/L_V &= (127 \pm 23) h_{70} M_{\odot} / L_{\odot, V} \\ (M/L_V &= (130 \pm 25) h_{70} M_{\odot} / L_{\odot, V} \quad \text{total sample}) , \end{aligned} \quad (7.24)$$

7.5. Results and discussion: mass-to-light ratio and mass-richness relation

which agrees well with our results for both cluster lists of $M/L_B = (115 \pm 13) h_{70} M_\odot/L_{\odot,B}$ (optical) and $M/L_B = (146 \pm 21) h_{70} M_\odot/L_{\odot,B}$ (X-ray). This result is also consistent with those of the individual redshift bins within the redshift range of Bahcall & Comerford (2002). Furthermore, it is very close to the value obtained by Hradecky et al. (2000) of $M/L_V \sim 140 h_{70} M_\odot/L_{\odot,V}$, for 8 nearby groups and clusters at $z \leq 0.09$, as well as being in agreement with that of Parker et al. (2005).

Muzzin et al. (2007) examine the relations between mass, luminosity, and temperature of 15 clusters with redshifts $0.19 < z < 0.55$. The masses in the study, M_{200} , are all virial masses. Each cluster's luminosity, $L_{200,K}$, is the total K -band luminosity measured within r_{200} of the cluster. Their results are

$$\frac{M_{200}}{M_\odot} = 10^{-0.95 \pm 2.21} \left(\frac{L_{200,K}}{L_{\odot,K}} \right)^{1.20 \pm 0.16} h_{70}^{-1.40 \pm 0.32}, \quad (7.25)$$

consistent with the slope obtained by Girardi et al. (2002). Overplotting onto Fig. 7.5 (green dashed line) we see that this relation also fits well with our data, however slightly worse than that of Girardi et al. (2002). Although the slope of the $M - L$ relation is expected to be the same in both K and B bands (see page 120), the use of different bands might be the reason for this slight difference.

Looking at how $M/L_{200,K}$ evolves with mass, Muzzin et al. (2007) compute

$$\frac{M_{200}}{L_{200,K}} = 10^{-6.92 \pm 2.04} \left(\frac{M_{200}}{h_{70} M_\odot} \right)^{0.57 \pm 0.13} h_{70} L_{\odot,K}/M_\odot. \quad (7.26)$$

The slope here is in good agreement with the results in our second redshift bin ($0.4 \leq z < 0.6$),

$$\begin{aligned} \frac{M}{L_B} &= 10^{1.73 \pm 0.49} \left(\frac{M}{h_{70} M_\odot} \right)^{0.47 \pm 0.54} h_{70} L_{\odot,K}/M_\odot \quad (\text{X-ray}) \\ \frac{M}{L_B} &= 10^{1.65 \pm 0.39} \left(\frac{M}{h_{70} M_\odot} \right)^{0.58 \pm 0.49} h_{70} L_{\odot,K}/M_\odot \quad (\text{optical}), \end{aligned} \quad (7.27)$$

however it does not agree with the much lower value of Popesso et al. (2007). This might point towards an increase of slope with redshift. Finally the mean $M/L_{200,K}$ for the Muzzin et al. (2007) clusters is estimated to $(61.2 \pm 1.8) h_{70} M_\odot/L_{\odot,K}$. This value is fairly low compared to the studies presented above, however, is in good agreement with the value calculated from our optical clusters in the second redshift bin.

7.5.4 The mass-richness relation

Given the tight correlation between luminosity and richness (see Sect. 7.5.2), we have only looked at the relation between mass and luminosity so far. However, it might be interesting to investigate the mass-richness relation in redshift bins in more detail.

Figure 7.9 shows the relation between mass and richness in redshift bins. Comparing it with the corresponding mass vs. light plot, Fig. 7.6, we see that, as expected, the two figures do not differ much. The results when fitting this relation to the whole data set yield

$$M/N_{\text{gal}} = (251 \pm 28) \times 10^{10} h_{70}^{-1} M_\odot \quad (\text{optical}) \quad (7.28)$$

$$M/N_{\text{gal}} = (288 \pm 38) \times 10^{10} h_{70}^{-1} M_\odot \quad (\text{X-ray}). \quad (7.29)$$

Mass-richness relation for all clusters in redshift bins

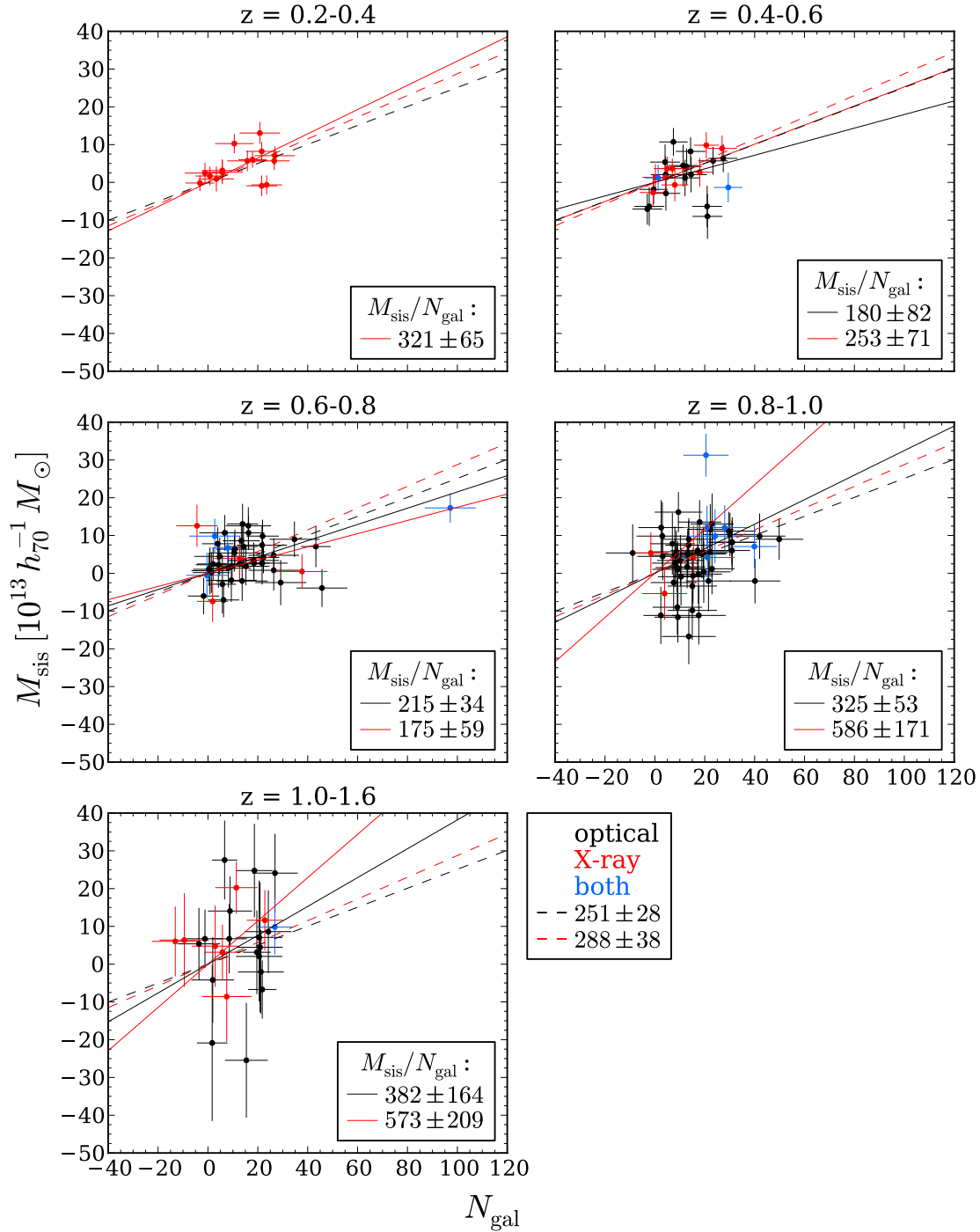


Figure 7.9: Mass as a function of richness for all clusters, divided in redshift bins. The mass-richness relations for each bin are drawn as solid lines, together with fits to the total datasets shown as dashed lines. The red points represent the X-ray clusters and the black points the optical clusters. The blue points are clusters present in both cluster catalogues, and are hence included in both fits. Note that $M_{\text{sis}}/N_{\text{gal}}$ is given in units of $[10^{10} h_{70}^{-1} M_{\odot}]$.

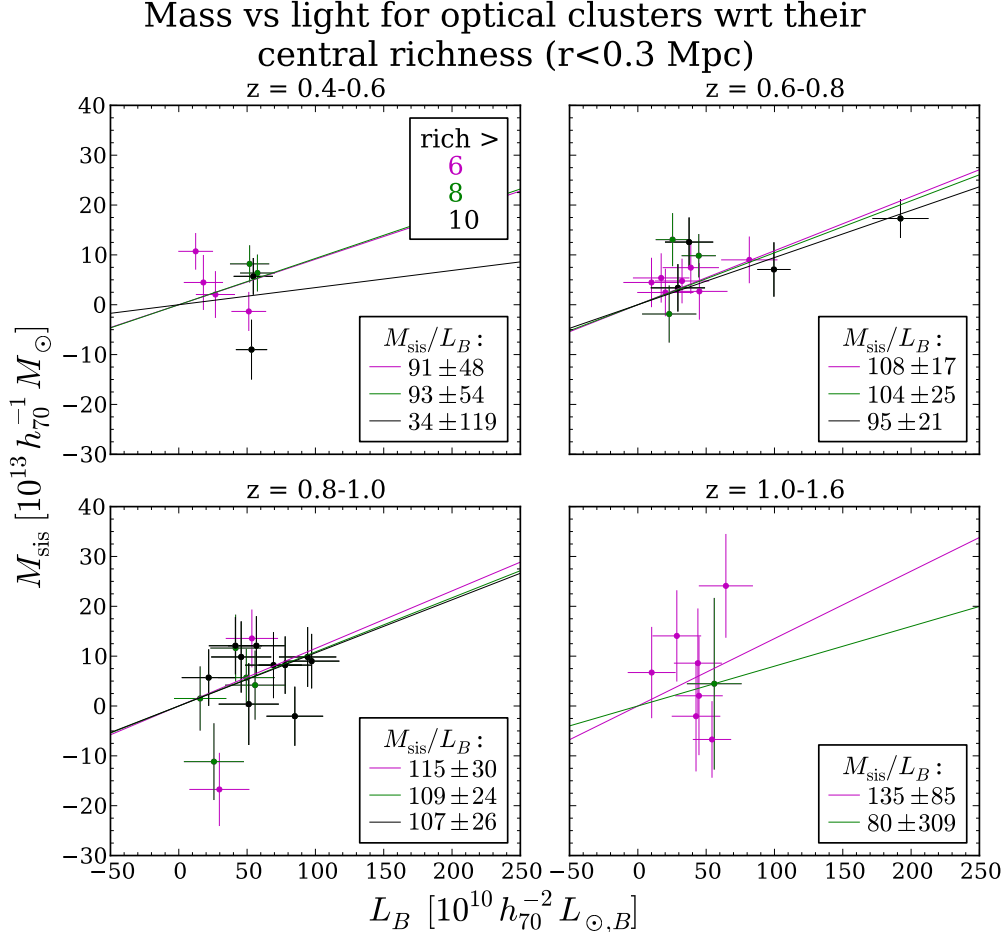


Figure 7.10: Mass as a function of optical luminosity for optical clusters with respect to their central richness ($r < 0.3$ Mpc), where the mass and luminosity have been calculated for $r = 1.0$ Mpc. Each richness bin includes the clusters with richnesses higher than the given limit, i.e. clusters with a central richness higher than 10 are black, those with a central richness higher than 8 are green and black, and all clusters in these plots have a central richness higher than 6. The mass-to-light ratios have been fit in separate redshift bins. Note that the green line in the last redshift bin has been forced through (0, 0) and the one data point in question.

7.5.5 Mass-to-light as a function of central richness

Investigating the densest clusters in greater detail can reduce some of the noise from lower-density clusters and yield more significant results for our optical clusters. In Fig. 7.10 we plot the mass-to-light relation for clusters sorted in redshift bins, applying different cuts with respect to central richness, where we define the central richness of a cluster as its richness within a central radius of 0.3 Mpc. We then fit a M/L_B ratio to the clusters having a central richness ≥ 10 , 8, and 6. Comparing the resulting M/L_B for the three thresholds, they are all consistent within the error bars. The three slopes seem to consistently increase with redshift, but the significance of this trend is low given the large fit uncertainties. One might also expect higher L_B at a given mass for richer systems, however, the number statistics are again too poor to draw any conclusions.

7.6 Results and discussion: comparing X-ray properties with optical measurements

In the previous section we looked at the relationship between the cluster mass inferred from gravitational lensing methods and their total optical luminosity. In this section we will concentrate on cluster properties derived by F07 from X-ray observations with *XMM-Newton* (see Sect. 7.2.2), by comparing these to the optical properties described earlier.

First, to understand more about how the measured cluster properties between the two widely different methods correlate, the X-ray cluster luminosity, $L_{0.1-2.4\text{keV}}$, is plotted as a function of the optical luminosity, L_B , in Fig. 7.11. In the redshift bins $0.4 \leq z < 0.6$, $0.6 \leq z < 0.8$, and $1.0 \leq z < 1.26$ we see some correlations between $L_{0.1-2.4\text{keV}}$ and L_B , although with a very large scatter. However, as the sample size is small it is difficult to draw statistically valid conclusions.

Figure 7.12 shows the lensing mass, M_{sis} , as a function of X-ray luminosity. The first two redshift bins, and to a certain degree the third, show a connection between the two quantities. The last two bins show no clear trend, however, performing lensing analyses at such high redshifts is difficult, which is also reflected in the M_{sis} error bars. Within error bars we see that $L_{0.1-2.4\text{keV}}$ is increasing with redshift for a given mass, however no statistically significant conclusion can be made. Comparing this figure to the $M_{\text{sis}} - L_B$ relation in Fig. 7.6 does not reveal any correlation between the two.

There is in general a tighter relation between a cluster's temperature and mass, than between its luminosity and mass (e.g. Pratt et al. 2009). So although the cluster X-ray temperatures in this chapter have been derived from the corresponding X-ray luminosities (Sect. 7.2.2), we still investigate the $M_{\text{sis}} - T_X$ relation in addition. The left panel of Fig. 7.13 shows M_{sis} as a function of T_X . Since this figure is very similar to Fig. 7.12, the whole sample is plotted in one figure and redshifts indicated with different colours. As for the $M_{\text{sis}} - L_{0.1-2.4\text{keV}}$ relation, the first three redshift bins show reasonable correlation, whereas the last two bins resemble no clear trends. Similarly, the right panel of Fig. 7.13 shows the relation between the lensing mass and the X-ray mass, M_{500} .

Measuring the optical luminosity, L , of a cluster is observationally much cheaper than measuring X-ray properties (e.g. Muzzin et al. 2007). If L can be used as a proxy for X-ray properties such as $L_{0.1-2.4\text{keV}}$, T_X , and M_{500} , follow-up observations of candidates in cluster cosmology projects are much easier to achieve. On the other hand, L has a large scatter with mass.

Figure 7.14 investigates how T_X is tracing L_B , where the different colours indicate different redshift ranges. We fit a power-law to the data set as a whole for T_X vs. L_B . We also fit $T_X(L_B)$ in redshift bins, but did not see a trend in $T_X \propto L_B$ with redshift. Due to small number statistics and the noisy data set we do not consider those results reliable enough to show here. Note also that due to the nature of the power-law, the input variable cannot be negative. Only clusters with $L_B > 0$ are therefore included in the fit. The result from fitting to the complete data set yields

$$k_B T_X = 10^{-0.96 \pm 0.23} \left(\frac{L_B h_{70}^2}{L_{\odot, B}} \right)^{0.73 \pm 0.13} \text{ keV} . \quad (7.30)$$

Looking at similar correlations, Muzzin et al. (2007) investigate whether L_K can be used

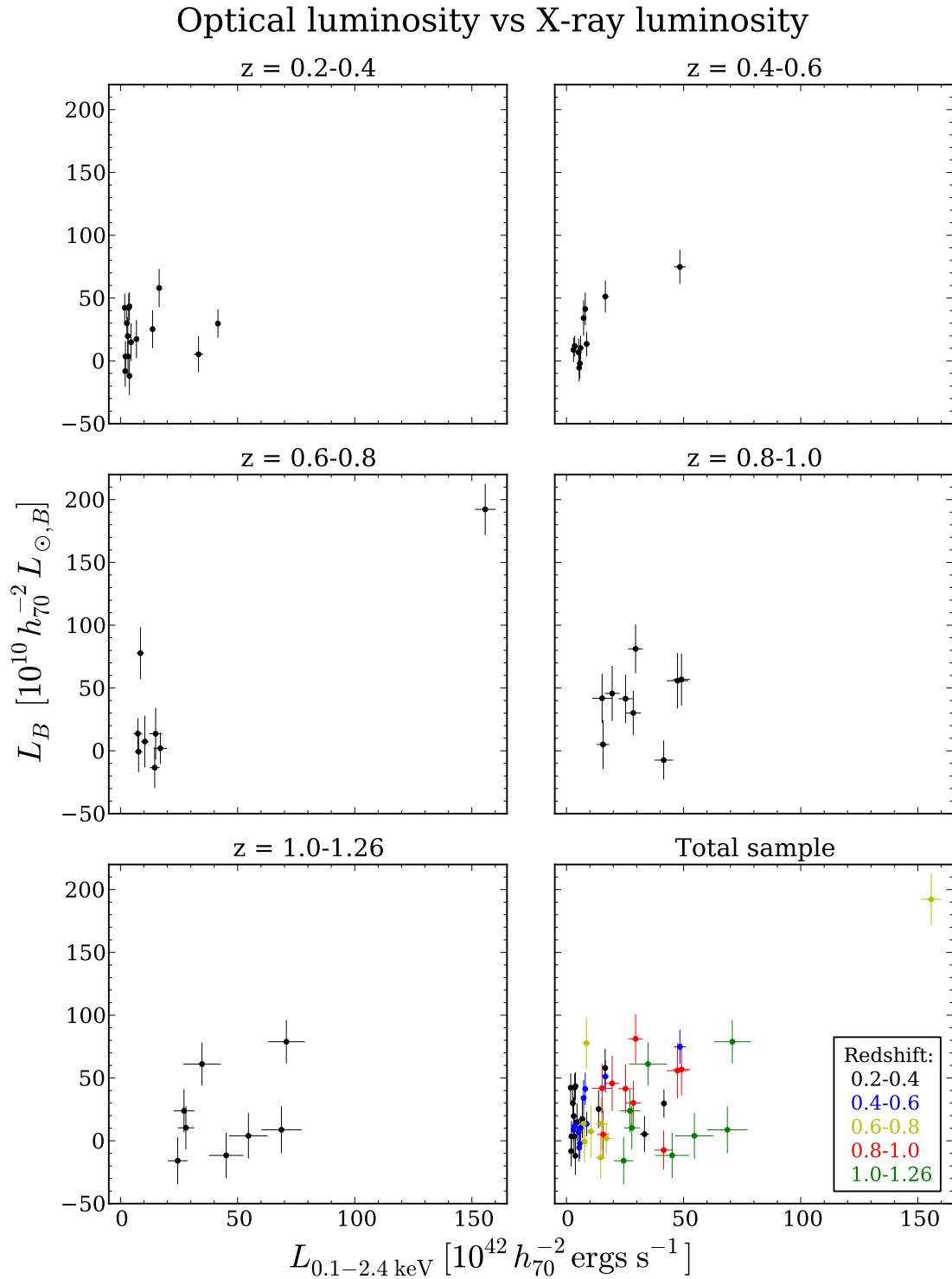


Figure 7.11: Optical vs. X-ray luminosity in redshift bins.

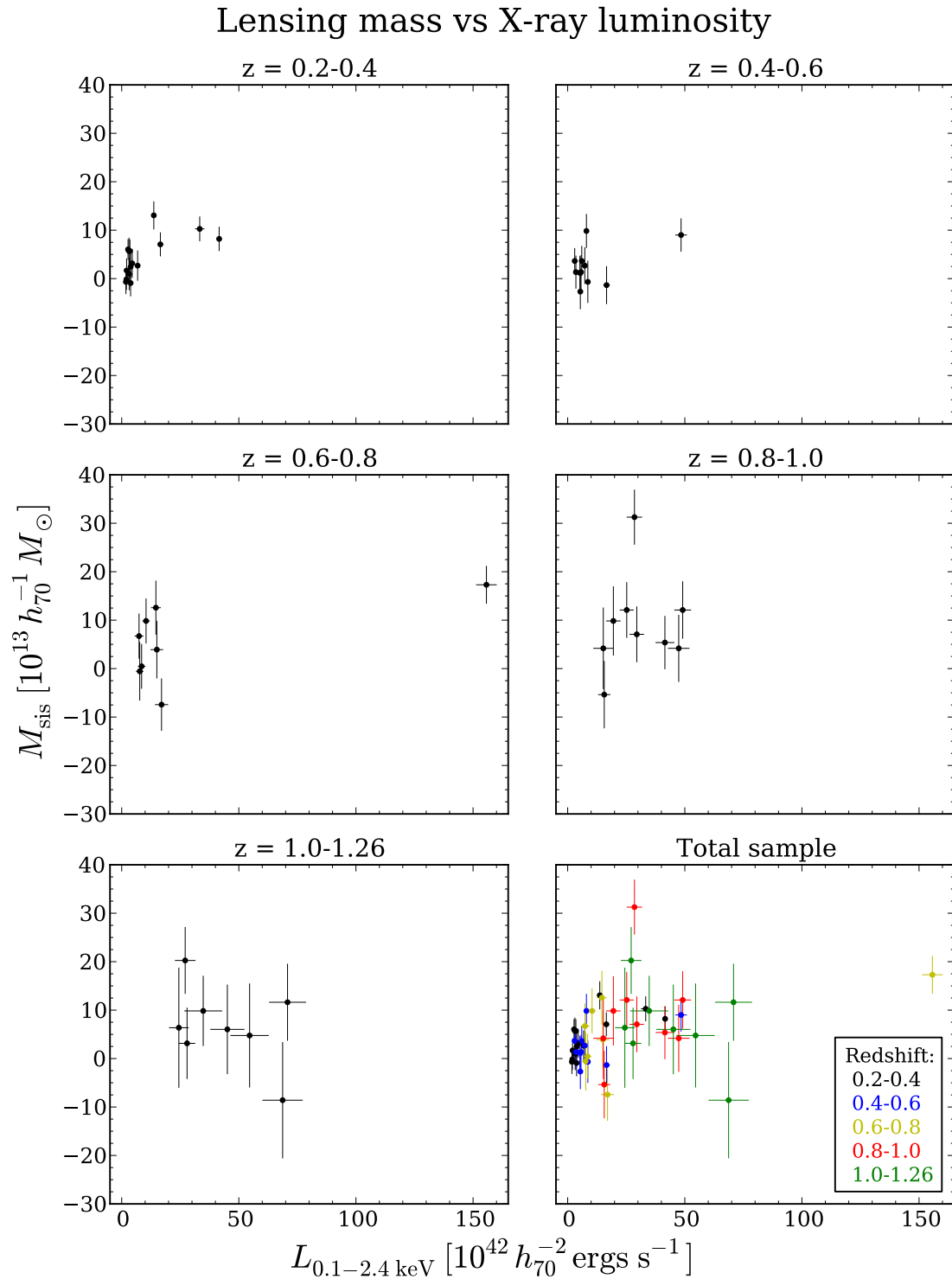


Figure 7.12: Lensing mass vs. X-ray luminosity in redshift bins. We see in the plot of the total sample that, for a given mass, $L_{0.1-2.4 \text{ keV}}$ is increasing with redshift.

7.6. Results and discussion: comparing X-ray and optical measurements

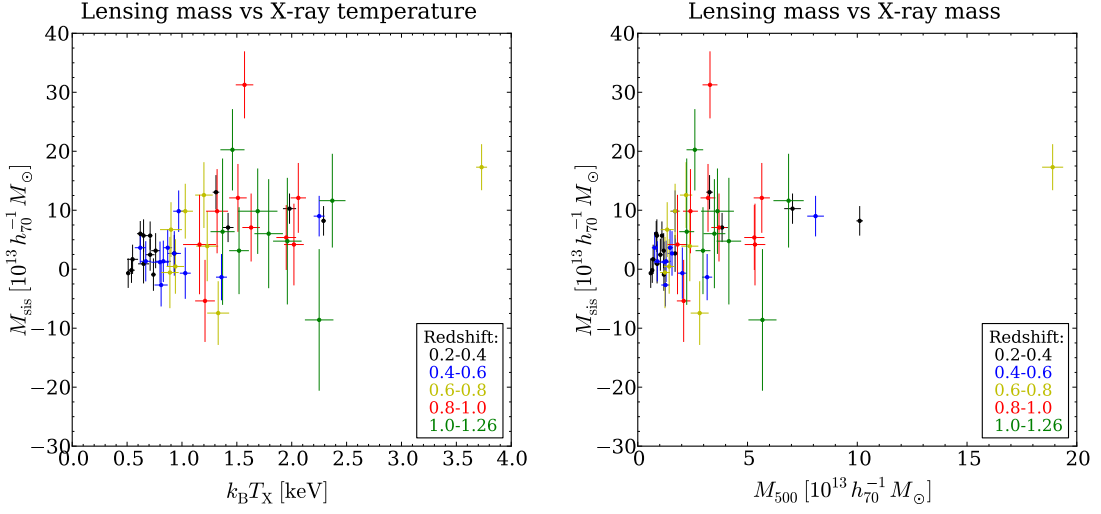


Figure 7.13: Lensing mass vs. X-ray temperature (*left*) and X-ray mass (*right*), both estimated from X-ray luminosity. Colours indicate redshift bins.

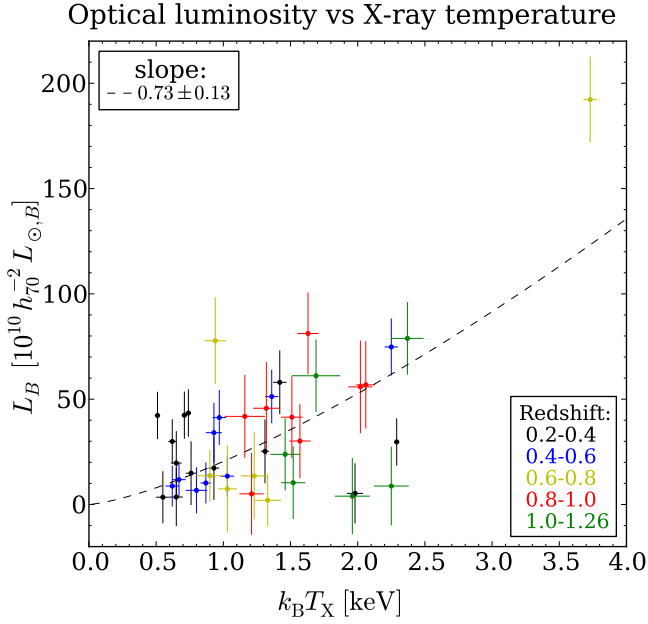


Figure 7.14: Optical luminosity vs. X-ray temperature for X-ray clusters. The temperature has been fitted to a power-law as a function of luminosity, $T_X = 10^\alpha L_B^{\text{slope}}$, represented by the solid line. The slope is given in the legend.

as a proxy for a cluster's X-ray luminosity, temperature, or mass. They find

$$k_B T_X = 10^{-9.45 \pm 1.07} \left(\frac{L_{200,K} h_{70}^2}{L_{\odot,K}} \right)^{0.77 \pm 0.08} \text{ keV}, \quad (7.31)$$

where the slope agrees very well with our best-fit slope.

7.7 Summary and conclusions

In this study we have analysed 178 galaxy clusters detected with two very different methods, one based on the spatial extent of X-ray emission and the other on galaxy overdensities in optical observations. We have investigated the mass properties of these clusters with weak gravitational lensing, and used optical data to examine their individual richness and luminosity. We have attempted to constrain the evolution of the cluster mass-to-light ratio as a function of redshift over a wide redshift range ($0.2 \leq z_{\text{cl}} < 1.6$). After correcting the cluster luminosities for passive evolution, we find no evolution in the M/L ratio with redshift for the X-ray clusters. For the optically selected clusters we see some indications for an increasing M/L when moving towards higher redshift. However, given the large uncertainties this is a low-significance detection, likely caused by either selection effects or spurious detections in the cluster catalogue.

With the exception of the evolution of M/L as a function of redshift, exploring the differences between the two datasets we find they generally agree very well. This applies to both redshift independent fits as well as when the samples have been divided into redshift bins. The results from this study are also in good agreement with the literature.

Appendix A

Cluster tables for Chapter 7

Table A.1: Optically selected cluster list by Zatloukal (2008). Clusters also found in the X-ray list (Table A.2) are labelled as *Common* and the X-ray cluster ID is given. Note that RA, DEC, and z for these clusters are taken from the X-ray cluster list (Sect. 7.2.3). An asterisk next to the cluster ID number indicates that the cluster has been removed from the final cluster sample (Sect. 7.2.4), the last three columns show under which criteria as explained in the following. *Outside field*: Clusters lying outside the COSMOS field, defined by the area covered by the Ilbert et al. (2009) photo- z catalogue. Note that these clusters have not been analysed, hence no information regarding shear or photo- z coverage has been obtained. *Shear*: Clusters with less than 50% spatial coverage in the Schrabback et al. (2010) shear catalogue. *Photo- z* : Clusters with less than 90% spatial coverage in the Ilbert et al. (2009) photo- z catalogue.

ID	RA [deg]	DEC [deg]	z	σ_z	Common	Outside field	Shear	Photo- z
1*	150.144	1.565	0.50	0.01			•	
2	150.576	2.163	0.51	0.04				
3	149.758	1.979	0.51	0.02	119			
4	150.215	1.825	0.51	0.01	65			
5*	150.327	1.504	0.51	0.01			•	
6*	150.192	1.496	0.52	0.01		•	–	–
7	150.361	1.633	0.52	0.04				
8	150.007	1.606	0.52	0.02				
9	149.761	1.605	0.52	0.03				
10	150.293	1.684	0.52	0.01				
11	150.131	1.857	0.53	0.01				
12*	150.059	1.518	0.53	0.03			•	•
13	150.048	1.626	0.53	0.01				
14	149.817	1.815	0.53	0.01				
15	149.665	1.887	0.55	0.03				
16	149.491	1.794	0.55	0.01				
17	149.518	1.878	0.56	0.01				
18*	149.418	1.849	0.56	0.01			•	•
19*	149.434	1.704	0.57	0.01			•	•
20*	149.806	1.580	0.57	0.01			•	

Appendix A

Table A.1: *Optically selected cluster list continued...*

ID	RA [deg]	DEC [deg]	z	σ_z	Common	Outside field	Shear	Photo- z
21	149.560	1.710	0.58	0.01				
22	149.625	1.735	0.58	0.01				
23	149.727	1.745	0.58	0.05				
24	149.904	1.686	0.59	0.03				
25*	149.680	1.549	0.59	0.01			•	
26*	149.596	1.517	0.59	0.02			•	•
27	149.571	1.782	0.59	0.01				
28*	149.522	1.498	0.60	0.01		•	–	–
29	150.486	2.755	0.60	0.01	36			
30	149.491	1.633	0.61	0.05				
31	149.800	1.722	0.62	0.01				
32	149.767	1.904	0.62	0.03				
33	150.256	1.974	0.62	0.01				
34*	150.150	2.910	0.62	0.02			•	•
35	149.608	1.890	0.62	0.01				
36	149.762	2.027	0.63	0.03				
37	150.196	2.157	0.64	0.01				
38*	150.277	1.583	0.64	0.03			•	
39	150.455	1.893	0.65	0.03	40			
40	150.059	2.800	0.65	0.06				
41	150.639	2.716	0.65	0.03				
42*	150.751	1.524	0.66	0.01			•	•
43	149.720	1.845	0.66	0.01				
44*	150.780	2.805	0.66	0.03			•	
45	150.502	2.447	0.66	0.01				
46	150.710	2.762	0.67	0.05				
47	149.948	2.098	0.67	0.01				
48	150.090	2.195	0.68	0.01				
49	150.163	2.506	0.68	0.01	73			
50	150.099	2.269	0.68	0.06				
51	150.054	2.311	0.68	0.01				
52	150.050	2.452	0.69	0.02				
53	149.654	2.825	0.69	0.02				
54*	150.272	2.535	0.69	0.02				•
55	150.141	2.067	0.69	0.01				
56	149.967	2.662	0.69	0.01				
57	149.960	2.541	0.70	0.01				
58*	150.062	2.413	0.70	0.01				•
59	149.843	2.400	0.70	0.01				
60	149.836	1.681	0.70	0.03				
61*	150.286	2.386	0.70	0.01				•
62	150.027	2.354	0.70	0.03				
63	149.785	2.465	0.70	0.02				

Table A.1: *Optically selected cluster list continued...*

ID	RA [deg]	DEC [deg]	z	σ_z	Common	Outside field	Shear	Photo-z
64	150.088	2.460	0.70	0.01				
65	149.899	2.393	0.71	0.01				
66	150.107	2.719	0.71	0.01				
67	150.147	2.603	0.71	0.01				
68*	150.242	1.495	0.72	0.05		•	–	–
69	150.594	2.127	0.72	0.04				
70*	150.317	2.871	0.72	0.02			•	
71	150.745	2.408	0.73	0.01				
72	149.678	2.264	0.74	0.01				
73*	149.405	2.453	0.74	0.03		•	–	–
74	149.552	2.009	0.76	0.01				
75	149.667	2.372	0.76	0.03				
76*	150.806	2.409	0.77	0.01			•	•
77	150.403	2.776	0.79	0.01				
78	150.663	2.818	0.79	0.01				
79*	150.829	2.467	0.80	0.02		•	–	–
80	150.214	2.853	0.80	0.03				
81	149.916	2.509	0.81	0.01	106			
82	150.469	2.554	0.81	0.02				
83*	150.299	2.481	0.81	0.01				•
84	150.526	2.139	0.81	0.01				
85	150.702	2.422	0.82	0.02				
86	149.718	2.709	0.83	0.02				
87	150.345	2.843	0.83	0.01				
88*	150.734	2.666	0.83	0.01				•
89	149.623	2.403	0.83	0.01				
90*	149.431	2.190	0.84	0.04			•	•
91*	149.610	2.860	0.84	0.07	134			•
92*	150.533	2.205	0.84	0.01	32			•
93*	150.065	2.407	0.84	0.05				•
94	150.365	2.004	0.84	0.01	53			
95	150.550	2.590	0.84	0.02				
96	150.431	1.965	0.85	0.01				
97	150.514	1.992	0.85	0.02				
98	150.091	2.527	0.86	0.01				
99*	150.588	2.871	0.86	0.01	24		•	
100	150.494	2.439	0.86	0.05				
101	150.448	2.140	0.87	0.01				
102	150.040	2.552	0.87	0.01				
103	150.181	2.588	0.87	0.01	72			
104	150.073	2.636	0.87	0.01				
105	150.666	2.380	0.87	0.01				
106	150.212	2.286	0.88	0.01	62			

Appendix A

Table A.1: *Optically selected cluster list continued...*

ID	RA [deg]	DEC [deg]	z	σ_z	Common	Outside field	Shear	Photo- z
107	150.506	2.569	0.88	0.03				
108	150.395	2.707	0.88	0.01				
109	149.958	2.664	0.88	0.01				
110*	149.434	1.965	0.88	0.01			•	•
111*	150.207	2.812	0.88	0.01				•
112	149.652	2.363	0.88	0.01				
113	149.922	2.631	0.88	0.01				
114	150.268	2.086	0.89	0.04				
115	150.419	2.515	0.89	0.01				
116	150.148	2.197	0.89	0.01				
117	149.966	2.349	0.89	0.01				
118	150.210	2.399	0.89	0.01	68			
119	150.430	1.851	0.89	0.01	45			
120	150.258	1.894	0.89	0.03				
121*	149.401	2.412	0.90	0.01		•	–	–
122*	150.259	2.899	0.90	0.01			•	•
123	150.149	2.795	0.90	0.03				
124	150.094	2.201	0.90	0.01	84			
125*	150.320	2.883	0.90	0.01			•	
126	149.909	2.702	0.90	0.01				
127	149.509	2.262	0.90	0.07				
128	150.025	2.203	0.91	0.01				
129	149.457	2.591	0.91	0.02				
130	149.493	2.014	0.91	0.01				
131	150.230	1.752	0.92	0.02				
132	149.843	2.574	0.92	0.01				
133	149.845	2.144	0.92	0.05				
134	149.660	2.234	0.93	0.01	128			
135	150.114	2.129	0.93	0.02				
136	150.321	2.280	0.94	0.02				
137	149.752	2.279	0.94	0.01				
138*	150.750	2.469	0.94	0.01			•	
139	149.898	2.335	0.94	0.01				
140	150.006	2.154	0.94	0.01				
141	150.161	2.686	0.95	0.05				
142	150.128	2.716	0.95	0.01				
143	150.077	2.684	0.96	0.02				
144	150.081	2.733	0.96	0.04				
145*	150.806	2.202	0.96	0.01			•	•
146	149.974	1.662	0.97	0.01				
147*	149.428	1.883	0.98	0.04			•	•
148*	149.739	2.351	0.98	0.02				•
149*	149.483	2.221	0.98	0.01				•

Table A.1: *Optically selected cluster list continued...*

ID	RA [deg]	DEC [deg]	z	σ_z	Common	Outside field	Shear	Photo-z
150*	150.807	2.688	0.99	0.02			•	•
151*	149.973	1.560	0.99	0.04			•	
152*	149.949	1.507	1.00	0.01			•	
153	150.121	2.002	1.00	0.05				
154	149.762	2.283	1.03	0.06				
155	150.685	2.285	1.04	0.06				
156	150.703	2.347	1.06	0.01				
157	150.568	2.497	1.08	0.04				
158	150.080	2.042	1.10	0.06				
159	150.439	2.751	1.12	0.08				
160*	150.606	1.751	1.17	0.06				•
161	150.585	2.094	1.20	0.06				
162	150.148	2.061	1.22	0.07				
163*	150.551	1.816	1.24	0.05				•
164*	150.227	2.784	1.26	0.07				•
165	150.099	2.699	1.27	0.03				
166	150.589	2.180	1.32	0.07				
167	149.990	2.691	1.32	0.07				
168	149.958	2.336	1.43	0.02				
169	149.976	2.490	1.45	0.06				
170	149.911	2.327	1.51	0.04				
171	149.910	2.358	1.55	0.07				
172	149.773	2.455	1.59	0.05				

Table A.2: X-ray selected cluster list by Finoguenov et al. (2007) (excluding 7 clusters that have $z < 0.2$, see the text for details). The last three columns indicate whether the cluster has been removed from the final cluster sample, see the Table A.1 caption for more information.

ID	RA [deg]	DEC [deg]	z	Common	Outside field	Shear	Photo-z
3*	150.80244	1.98985	0.25			•	•
9*	150.75121	1.52793	0.75			•	•
11*	150.73676	2.82680	0.60			•	
15	150.67342	2.09190	0.34				
20*	150.64041	2.12791	0.55				•
24*	150.58962	2.87187	0.95	99		•	
25*	150.58631	1.92693	0.30				•
32*	150.50535	2.22395	0.90	92			•
34	150.49330	2.06795	0.40				
36	150.49048	2.74592	0.65	29			
38	150.44824	1.91197	1.25				
39	150.44827	2.04996	0.55				
40	150.44523	1.88197	0.70	39			

Appendix A

Table A.2: *X-ray cluster list continued...*

ID	RA [deg]	DEC [deg]	z	Common	Outside field	Shear	Photo-z
41	150.44229	2.15796	0.40				
44	150.42124	1.98397	0.45				
45	150.42121	1.84898	0.85	119			
47	150.40934	2.51196	1.00				
51	150.37616	1.66900	0.75				
52	150.37930	2.40997	0.35				
53	150.37021	1.99898	0.85	94			
54	150.33413	1.60301	0.40				
56	150.31318	2.00799	0.35				
57*	150.28611	1.55502	0.36			•	
59*	150.28311	1.57902	0.36			•	
62	150.21114	2.28100	0.88	106			
64*	150.23218	2.48199	0.30				•
65	150.21111	1.81600	0.53	4			
66	150.21718	2.73998	0.95				
67	150.19609	1.65701	0.22				
68	150.21115	2.40100	0.90	118			
70	150.18109	1.76801	0.35				
71	150.19616	2.82397	0.20				
72	150.16011	2.60499	0.90	103			
73	150.16912	2.52400	0.75	49			
75	150.15410	2.39500	1.15				
78	150.11807	2.35600	0.22				
79	150.11807	2.68299	0.35				
80	150.10906	2.55700	0.50				
82	150.10606	2.42200	0.22				
83	150.10906	2.01400	0.85				
84	150.09405	2.20000	0.93	124			
85	150.09105	2.39500	0.22				
86	150.09705	2.30200	0.36				
87*	150.05802	2.38000	0.40				•
89	150.03999	2.69499	0.20				
93	150.04300	2.54500	1.25				
97	149.98594	2.58099	0.70				
99	149.96500	1.68101	0.37				
100	149.97091	2.78197	0.70				
101	149.96495	2.21199	0.43				
102	149.95293	2.34099	1.10				
103	149.94991	2.48199	0.80				
105	149.91987	2.60198	0.25				
106	149.91688	2.51498	0.73	81			
108	149.88981	2.80596	0.65				
111	149.88386	2.44898	0.36				

Table A.2: *X-ray cluster list continued...*

ID	RA [deg]	DEC [deg]	z	Common	Outside field	Shear	Photo-z
114	149.81183	2.25397	0.47				
119	149.74586	1.94797	0.45	3			
120	149.75467	2.79393	0.49				
126	149.64969	2.34093	1.00				
128	149.64372	2.21193	1.00	134			
132*	149.59548	2.82087	0.34				•
133	149.60532	2.43541	1.15				
134*	149.60148	2.85087	0.95	91			•
145*	149.39739	2.5748	0.37		•	–	–

Appendix A

Bibliography

- Abate, A., Wittman, D., Margoniner, V. E., et al. 2009, ApJ, 702, 603
- Abell, G. O. 1958, ApJS, 3, 211
- Adami, C., Mazure, A., Biviano, A., Katgert, P., & Rhee, G. 1998a, A&A, 331, 493
- Adami, C., Mazure, A., Katgert, P., & Biviano, A. 1998b, A&A, 336, 63
- Allen, S. W., Rapetti, D. A., Schmidt, R. W., et al. 2008, MNRAS, 383, 879
- Allen, S. W., Schmidt, R. W., Fabian, A. C., & Ebeling, H. 2003, MNRAS, 342, 287
- Amanullah, R., Lidman, C., Rubin, D., et al. 2010, ApJ, 716, 712
- Araya-Melo, P. A., Reisenegger, A., Meza, A., et al. 2009, MNRAS, 399, 97
- Bahcall, J. N. & Sarazin, C. L. 1977, ApJ, 213, L99
- Bahcall, N. A. & Comerford, J. M. 2002, ApJ, 565, L5
- Bartelmann, M. 1996, A&A, 313, 697
- Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291
- Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393
- Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517
- Böhringer, H., Schuecker, P., Guzzo, L., et al. 2004, A&A, 425, 367
- Borgani, S. & Guzzo, L. 2001, Nature, 409, 39
- Broadhurst, T., Takada, M., Umetsu, K., et al. 2005, ApJ, 619, L143
- Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559
- Carlberg, R. G., Yee, H. K. C., Ellingson, E., et al. 1996, ApJ, 462, 32
- Carlberg, R. G., Yee, H. K. C., Morris, S. L., et al. 1999, ApJ, 516, 552
- Carlberg, R. G., Yee, H. K. C., Morris, S. L., et al. 2001, ApJ, 552, 427
- Cassano, R. & Brunetti, G. 2005, MNRAS, 357, 1313
- Chandrasekhar, S. 1943, ApJ, 97, 255
- Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJ, 648, L109
- Clowe, D., De Lucia, G., & King, L. 2004, MNRAS, 350, 1038
- Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, AJ, 132, 926
- Comerford, J. M. & Natarajan, P. 2007, MNRAS, 379, 190
- Comins, N. F. & Kaufmann, W. J. 2003, *Discovering the Universe*, W. H. Freeman
- Corless, V. L. & King, L. J. 2007, MNRAS, 380, 149
- Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20
- Dahle, H., Kaiser, N., Irgens, R. J., Lilje, P. B., & Maddox, S. J. 2002, ApJS, 139, 313
- Danese, L., de Zotti, G., & di Tullio, G. 1980, A&A, 82, 322
- David, L. P., Jones, C., & Forman, W. 1995, ApJ, 445, 578

Bibliography

- Dietrich, J. P., Schneider, P., Clowe, D., Romano-Díaz, E., & Kerp, J. 2005, *A&A*, 440, 453
- Donahue, M., Voit, G. M., Gioia, I., et al. 1998, *ApJ*, 502, 550
- Dressler, A. 1978, *ApJ*, 226, 55
- Dressler, A. 1984a, *ApJ*, 281, 512
- Dressler, A. 1984b, *ARA&A*, 22, 185
- Einstein, A. 1915, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, Seite 844-847., 844
- Einstein, A. 1916, *Annalen der Physik*, 354, 769
- Ellis, R. S., Colless, M., Broadhurst, T., Heyl, J., & Glazebrook, K. 1996, *MNRAS*, 280, 235
- Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, *Astronomische Nachrichten*, 326, 432
- Evrard, A. E., MacFarland, T. J., Couchman, H. M. P., et al. 2002, *ApJ*, 573, 7
- Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, *ApJ*, 513, 34
- Finoguenov, A., Guzzo, L., Hasinger, G., et al. 2007, *ApJS*, 172, 182
- Finoguenov, A., Reiprich, T. H., & Böhringer, H. 2001, *A&A*, 368, 749
- Gavazzi, R. 2005, *A&A*, 443, 793
- Giacintucci, S., Venturi, T., Cassano, R., Dallacasa, D., & Brunetti, G. 2009, *ApJ*, 704, L54
- Giavalisco, M. 2002, *ARA&A*, 40, 579
- Girardi, M., Borgani, S., Giuricin, G., Mardirossian, F., & Mezzetti, M. 2000, *ApJ*, 530, 62
- Girardi, M., Manzato, P., Mezzetti, M., Giuricin, G., & Limboz, F. 2002, *ApJ*, 569, 720
- Gladders, M. & Yee, H. 2000, *AJ*, 120, 2148
- Harrison, E. R. 1974, *ApJ*, 191, L51
- Hartle, J. B. 2003, *Gravity: an introduction to Einstein's general relativity*, Addison Wesley
- Herschel, W. 1785, *Royal Society of London Philosophical Transactions Series I*, 75, 213
- Heymans, C., Gray, M. E., Peng, C. Y., et al. 2008, *MNRAS*, 385, 1431
- Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, *MNRAS*, 368, 1323
- Hildebrandt, H., Erben, T., Dietrich, J. P., et al. 2006, *A&A*, 452, 1121
- Hildebrandt, H., Pielorz, J., Erben, T., et al. 2009, *A&A*, 498, 725
- Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, *ApJ*, 504, 636
- Hoekstra, H., Franx, M., Kuijken, K., & van Dokkum, P. G. 2002, *MNRAS*, 333, 911
- Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J. 2002, *arXiv:astro-ph/0210394*
- Holhjem, K., Schirmer, M., & Dahle, H. 2009, *A&A*, 504, 1
- Howell, S. B. 2006, *Handbook of CCD astronomy*, Cambridge University Press
- Hradecky, V., Jones, C., Donnelly, R. H., et al. 2000, *ApJ*, 543, 521
- Hubble, E. 1929a, *Proceedings of the National Academy of Science*, 15, 168
- Hubble, E. P. 1929b, *Leaflet of the Astronomical Society of the Pacific*, 1, 93
- Hubble, E. P. 1929c, *ApJ*, 69, 103
- Hubble, E. P. 1936, *Realm of the Nebulae*, New Haven: Yale University Press
- Huchra, J. P. & Geller, M. J. 1982, *ApJ*, 257, 423
- Ilbert, O., Capak, P., Salvato, M., et al. 2009, *ApJ*, 690, 1236
- Irgens, R. J., Lilje, P. B., Dahle, H., & Maddox, S. J. 2002, *ApJ*, 579, 227

- Jee, M. J., Rosati, P., Ford, H. C., et al. 2009, *ApJ*, 704, 672
- Kaiser, N. 2000, *ApJ*, 537, 555
- Kaiser, N. & Squires, G. 1993, *ApJ*, 404, 441
- Kaiser, N., Squires, G., & Broadhurst, T. 1995, *ApJ*, 449, 460
- Kapteyn, J. C. 1922, *ApJ*, 55, 302
- Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, *MNRAS*, 303, 188
- King, L. & Corless, V. 2007, *MNRAS*, 374, L37
- King, L. J., Clowe, D. I., Lidman, C., et al. 2002, *A&A*, 385, L5
- Klimov, Y. G. 1963, *Soviet Phys. Doklady*, 8, 119
- Kneib, J.-P., Hudelot, P., Ellis, R. S., et al. 2003, *ApJ*, 598, 804
- Kochanek, C. S., White, M., Huchra, J., et al. 2003, *ApJ*, 585, 161
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2010, *ArXiv e-prints*
- Koo, D. C. & Kron, R. G. 1992, *ARA&A*, 30, 613
- Landolt, A. U. 1983, *AJ*, 88, 439
- Landolt, A. U. 1992, *AJ*, 104, 340
- Lehnert, M. D., Nesvadba, N. P. H., Cuby, J., et al. 2010, *Nature*, 467, 940
- Liebes Jr., S. 1964, *Phys. Rev.*, 133, B835
- Limousin, M., Richard, J., Jullo, E., et al. 2007, *ApJ*, 668, 643
- Lin, W. P., Jing, Y. P., Mao, S., Gao, L., & McCarthy, I. G. 2006, *ApJ*, 651, 636
- Lin, Y., Mohr, J. J., & Stanford, S. A. 2003, *ApJ*, 591, 749
- Lin, Y., Mohr, J. J., & Stanford, S. A. 2004, *ApJ*, 610, 745
- Luppino, G. A. & Kaiser, N. 1997, *ApJ*, 475, 20
- Mantz, A., Allen, S. W., Rapetti, D., & Ebeling, H. 2010, *MNRAS*, 406, 1759
- Markevitch, M. 1998, *ApJ*, 504, 27
- Massey, R., Heymans, C., Bergé, J., et al. 2007a, *MNRAS*, 376, 13
- Massey, R., Rhodes, J., Leauthaud, A., et al. 2007b, *ApJS*, 172, 239
- Mathews, W. G. 1978, *ApJ*, 219, 413
- Mattig, W. 1958, *Astronomische Nachrichten*, 284, 109
- Messier, C. 1784, *Connaissance des Temps* (Paris)
- Miralda-Escude, J. 1991, *ApJ*, 370, 1
- Mobasher, B., Capak, P., Scoville, N. Z., et al. 2007, *ApJS*, 172, 117
- Monet, D. G., Levine, S. E., Canzian, B., et al. 2003, *AJ*, 125, 984
- Munshi, D., Valageas, P., van Waerbeke, L., & Heavens, A. 2008, *Phys. Rep.*, 462, 67
- Muzzin, A., Yee, H. K. C., Hall, P. B., & Lin, H. 2007, *ApJ*, 663, 150
- Narayan, R. & Bartelmann, M. 1996, *arXiv:astro-ph/9606001v2*
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1995, *MNRAS*, 275, 720
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
- Oguri, M., Takada, M., Umetsu, K., & Broadhurst, T. 2005, *ApJ*, 632, 841
- Oort, J. H. 1958, in *La Structure et L'Évolution de L'Univers*, Onzième Conseil de Physique, ed. R. Stoops (Solvay: Bruxelles), 163
- Parker, L. C., Hudson, M. J., Carlberg, R. G., & Hoekstra, H. 2005, *ApJ*, 634, 806
- Patel, S. K., Joy, M., Carlstrom, J. E., et al. 2000, *ApJ*, 541, 37
- Peacock, J. A. 1999, *Cosmological Physics*, Cambridge University Press
- Pedersen, K. & Dahle, H. 2007, *ApJ*, 667, 26
- Pires, S., Starck, J., & Refregier, A. 2010, *IEEE Signal Processing Magazine*, 27, 76
- Popesso, P., Biviano, A., Böhringer, H., & Romaniello, M. 2007, *A&A*, 464, 451

Bibliography

- Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, *A&A*, 498, 361
- Ramella, M., Boschin, W., Geller, M. J., Mahdavi, A., & Rines, K. 2004, *AJ*, 128, 2022
- Refsdal, S. 1964a, *MNRAS*, 128, 307
- Refsdal, S. 1964b, *MNRAS*, 128, 295
- Reiprich, T. H. 2001, PhD thesis, Max-Planck-Institut für extraterrestrische Physik, Garching, Germany
- Riess, A. G., Macri, L., Casertano, S., et al. 2009, *ApJ*, 699, 539
- Rines, K., Geller, M. J., Diaferio, A., Kurtz, M. J., & Jarrett, T. H. 2004, *AJ*, 128, 1078
- Rood, H. J. & Sastry, G. N. 1971, *PASP*, 83, 313
- Röser, H., Hippelein, H., Wolf, C., Zatloukal, M., & Falter, S. 2010, *A&A*, 513, A15
- Schaeffer, R., Maurogordato, S., Cappi, A., & Bernardreau, F. 1993, *MNRAS*, 263, L21
- Schirmer, M., Erben, T., Hettterscheidt, M., & Schneider, P. 2007, *A&A*, 462, 875
- Schirmer, M., Erben, T., Schneider, P., et al. 2003, *A&A*, 407, 869
- Schneider, P. 1996, *MNRAS*, 283, 837
- Schneider, P. 2006a, in: *Gravitational Lensing: Strong, Weak & Micro*, Saas-Fee Advanced Course 33, Swiss Society for Astrophysics and Astronomy, G. Meylan, P. Jetzer & P. North (Eds.), Springer-Verlag: Berlin, p. 1
- Schneider, P. 2006b, in: *Gravitational Lensing: Strong, Weak & Micro*, Saas-Fee Advanced Course 33, Swiss Society for Astrophysics and Astronomy, G. Meylan, P. Jetzer & P. North (Eds.), Springer-Verlag: Berlin, p. 269
- Schneider, P. 2006c, *Extragalactic Astronomy and Cosmology*, Springer-Verlag: Berlin
- Schneider, P. 2008, *Cosmology*, Lecture Notes, Univ. of Bonn
- Schneider, P., King, L., & Erben, T. 2000, *A&A*, 353, 41
- Schneider, P. & Seitz, C. 1995, *A&A*, 294, 411
- Schrabback, T. 2007, PhD thesis, Univ. of Bonn, Germany
- Schrabback, T., Hartlap, J., Joachimi, B., et al. 2010, *A&A*, in press [arXiv:astro-ph/0911.0053]
- Scoville, N., Abraham, R. G., Aussel, H., et al. 2007a, *ApJS*, 172, 38
- Scoville, N., Aussel, H., Brusa, M., et al. 2007b, *ApJS*, 172, 1
- Seitz, C. & Schneider, P. 1995, *A&A*, 297, 287
- Seitz, C. & Schneider, P. 1997, *A&A*, 318, 687
- Seitz, S. & Schneider, P. 2001, *A&A*, 374, 740
- Sheldon, E. S., Johnston, D. E., Masjedi, M., et al. 2009, *ApJ*, 703, 2232
- Spitzer, Jr., L. & Baade, W. 1951, *ApJ*, 113, 413
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, *ApJ*, 462, L17
- Stetson, P. B. 2000, *PASP*, 112, 925
- Sunyaev, R. A. & Zel'dovich, Y. B. 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Suyu, S. H., Marshall, P. J., Auger, M. W., et al. 2010, *ApJ*, 711, 201
- van Dokkum, P. G. & Franx, M. 2001, *ApJ*, 553, 90
- Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., et al. 2009, *ApJ*, 692, 1060
- Voit, G. M. 2005, *Reviews of Modern Physics*, 77, 207
- Wambsganss, J. 2006, in: *Gravitational Lensing: Strong, Weak & Micro*, Saas-Fee Advanced Course 33, Swiss Society for Astrophysics and Astronomy, G. Meylan, P. Jetzer & P. North (Eds.), Springer-Verlag: Berlin, p. 453

Bibliography

- Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, *ApJ*, 646, 881
Wold, M., Lacy, M., Dahle, H., Lilje, P. B., & Ridgway, S. E. 2002, *MNRAS*, 335, 1017
Wright, C. O. & Brainerd, T. G. 2000, *ApJ*, 534, 34
Zatloukal, M. 2008, PhD thesis, Heidelberg University, Germany
Zatloukal, M., Röser, H., Wolf, C., Hippelein, H., & Falter, S. 2007, *A&A*, 474, L5
Zwicky, F. 1933, *Helv. Phys. Acta*, 6, 110
Zwicky, F. 1937, *ApJ*, 86, 217

Bibliography

Acknowledgements

First and foremost I would like to thank my wonderful husband, who has made my life enjoyable despite difficult PhD times. Thank you for your unending love, support, and understanding. Thank you also for untiringly reading and correcting countless (more or less finished) versions of this thesis, and for your thorough explanations of all things incomprehensible for my tired PhD brain. I also want to thank my parents. Thank you so much for all your support, concern, love, and prayers. Thank you for accepting and loving me regardless of my mood and stress level, and for trying to understand these incomprehensible circumstances I was going through. I would like to thank my brothers and sisters(-in-law), my Norwegian friends and extended family, as well as my friends at the APC for all your support, love, and prayers.

Special thanks go to Tim Schrabback for proposing and working with me on several of the projects presented in this thesis. Tim, I would particularly like to thank you for your warm friendship and for opening your home to me every time I visited Leiden. You made work so much easier! I would like to thank Peter Schneider and Thomas Erben for being my supervisors throughout this PhD. I also want to thank Peter and Thomas, along with Thomas Reiprich, for careful reading and correcting of various thesis drafts. In addition I would like to thank Reiner Vianden and Andreas Hense for being on my thesis committee. I am grateful to Henk Hoekstra for insightful comments as well as useful and constructive discussions regarding my projects. I would also like to thank Konrad Kuijken and Leiden University for your hospitality during my regular visits to Leiden. Warm thanks go to all members of CFHTLenS¹ and the Argelander-Institut für Astronomie, particularly those of the lensing group and the observational group, the secretaries, and my lovely office mates.

Finally I would like to thank the Research council of Norway for support from a doctoral fellowship, project number 177254/V30, and in particular my contact at the RCN, Bjørn Jacobsen.

¹cfhtlens.org

Acknowledgements

Declaration

I herewith declare that I prepared this doctoral thesis myself, unaided and independently, using the literature listed in the bibliography.

Karianne Holhjem
Bonn, February 14, 2011

Declaration

Curriculum Vitae

Personal information

Full name: Karianne Holhjem Schirmer
Address: Mühlenstrasse 22, D-53347 Oedekoven, Germany
Date and place of birth: December 16, 1980, Oslo, Norway
Citizenship: Norwegian
Marital status: Married
E-mail: kholhjem@astro.uni-bonn.de
Phone: +49 228 33610939 / +49 176 67238894

Professional working experience

Aug 2005 - Dec 2006 The Synnøve Irgens-Jensen Distinguished Research Studentship at the Nordic Optical Telescope (NOT), working as a student support astronomer, Santa Cruz de La Palma, Spain

Higher education

Jan 2007 - Feb 2011 PhD thesis in astronomy at AIfA, University of Bonn, Germany, supported by the Research Council of Norway grant 177254/V30
Jan 2004 - Apr 2006 Master of Science degree in astronomy, University of Oslo, Norway. The master degree contains courses in astrophysics, physics, programming, and statistics, together with a master thesis on weak gravitational lensing
Aug 2000 - Dec 2003 Bachelor of Science degree in physics, University of Bergen, Norway. The bachelor degree contains courses in physics, mathematics, programming, statistics, intercultural studies, and philosophical exams
Aug 1999 - Jun 2000 Bible School of Bildøy Bergen, Norway
Aug 1996 - Jun 1999 High school, the general branch study, Sandefjord VGS, Norway
Aug 1993 - Jun 1996 Gjøklep Udomsskole, Holmestrand, Norway
Aug 1987 - Jun 1993 Ekeberg Skole, Holmestrand, Norway

Curriculum Vitae